丘成桐主编

数学翻译丛书

高等微积分（修订版）

Advanced Calculus (Revised Edition)

■ Lynn H. Loomis Shlomo Sternberg

■ 王元 胥鸣伟 译

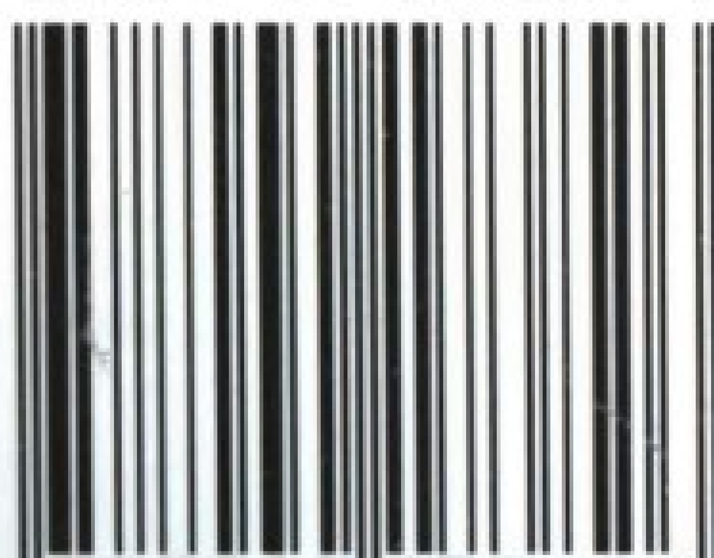


高等教育出版社

Higher Education Press

本书是哈佛大学的高等微积分教材，内容涵盖了从基本的向量空间概念到经典力学基本定理，包括多元微积分、外微分、微分形式的积分等。本书的特点是作者从拓扑－几何的观点来写微积分，用更现代的方式讲线性代数，把线性代数与微积分紧密地结合起来，这顺应了当代数学“拓扑几何与分析结合”的发展潮流。作者语言简练、优美、易懂，习题十分丰富而有价值。

本书可供数学类专业的本科高年级和研究生作为参考书使用，对广大教师和研究人员也有很好的参考价值。


ISBN 7-04-017382-4
9 787040 173826 >
定价 68.00 元

丘成桐主编
数学翻译丛书

高等微积分（修订版）

Advanced Calculus (Revised Edition)

■ Lynn H. Loomis　Shlomo Sternberg
■ 王元　胥鸣伟 译

高等教育出版社
Higher Education Press
International Press

图字: 01-2005-3463
ORIGINAL ENGLISH LANGUAGE EDITION PUBLISHED BY
Jones and Bartlett Publishers, Inc.
40 Tall Pine Drive
Sudbury, MA 01776

图书在版编目 (CIP) 数据

高等微积分 / (美) 卢米斯 (Loomis, L.H.), (美) 斯滕伯格 (Sternberg, S.) 王元, 胥鸣伟译. —修订本. —北京: 高等教育出版社, 2005.7
(数学翻译丛书 / 丘成桐主编)
书名原文: Advanced Calculus
ISBN 7-04-017382-4

I.高… II.①卢…②斯…③王…④胥…
III. 微积分-高等学校-教学参考资料 IV.0172

中国版本图书馆 CIP 数据核字 (2005) 第 061206 号

策划编辑	张小萍	责任编辑	赵天夫	封面设计	王凌波
版式设计	孙颖超	责任绘图	杜晓丹	责任印制	韩　刚

出版发行	高等教育出版社	购书热线	010-58581118
社　　址	北京市西城区德外大街4号	免费咨询	800-810-0598
邮政编码	100011	网　　址	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
经　　销	北京蓝色畅想图书发行有限公司	网上订购	http://www.landraco.com
印　　刷	北京原创阳光印业有限公司		http://www.landraco.com.cn
开　　本	787×960　1/16	版　　次	2005年7月第1版
印　　张	41.5	印　　次	2005年7月第1次印刷
字　　数	730 000	定　　价	68.00元

物 料 号　17382-00

原书前言

本书的基础是我们在 20 世纪 60 年代的一门高等微积分的荣誉课程. 在第一章到第十一章的那些不带星号的章节中包括了按常规所说的基本内容, 但是这些基本内容的不同应用被年复一年地加强, 从而使本书包含了比在任何一年可容纳的更多内容. 因此它可以被用作 (做一些删节) 高等微积分一年课程的教科书或者三学期的分析导引的教材.

必备的知识从严格的数学观点来说, 是对单变量微积分的良好基础知识连同对线性代数的某种程度的了解. 读者应该熟悉极限和连续性类型的论证方式, 并且有相当的数学修养. 至于可以提出的有关基础知识的教科书有 R. 柯朗 (Courant) 的《Differential and Integral Calculus》, T. 阿波斯图尔 (Apostol) 的《Calculus》, M. 斯皮瓦克 (Spivak) 的《Calculus》和 G. 哈代 (Hardy) 的《Pure Mathematics》. 读者还应该有求偏导数的经验.

按总的计划, 本书大体上分成上半部和下半部. 上半部讲述了在赋范向量空间架构上的微积分 (主要是微分), 而下半部处理的是微分流形上的微积分.

处理向量空间的微积分用了两章: 第三章的微分和第六章的常微分方程基本理论. 此前的那些章节是辅助性的. 第一章和第二章讲述了所需要的向量空间的纯代数理论. 第四章提供了关于紧性和完备性的内容, 这些是微积分的更为重要的结果所需要的. 第五章包含了在内积空间中所遇到的额外结构的一个简短阐述. 第七章致力于多重线性 (张量) 代数, 它主要是为以后要用到的参考章节.

第八章处理了欧氏空间上的 (黎曼) 积分理论并包括了关于傅里叶变换的基本事实 (以习题的形式). 第九章和第十章讲述了流形上的微积分, 而第十一章处理了 E. 嘉当的外微分和积分.

前面十一章形成一个逻辑单元, 其中每一章均依赖于前面各章节的结果. (当然, 许多章都包含有首次阅读时可以略去的材料; 它们一般出现在带有星号的章节中.) 另一方面, 第十二章和第十三章以及第六章和第十一章的后面部分是相互独立的, 它们被当作对前面各章所发展出的方法的带有解释性质的应用. 这里提出的有初等施图姆 - 刘维尔理论和傅里叶级数, 初等微分几何, 位势理论, 以及经典力学. 通常我们在一年的课程中只讲述这些专题中的一两个.

对于同一素材我们会从不同的观点毫不犹豫地多次提出讨论. 例如, 对于隐函数定理, 虽然我们已经选取了压缩映像不动点定理作为我们基本的处理方法, 然而我们也在正文中概述了一个 "牛顿法" 的证明, 并且在习题中也描述了第三种证明. 类似地, 变分法也遇到了两次: 一次在无穷维向量空间的微分学中, 而后则在经典力学的讨论中. 在最初几章中引进了向量空间的子流形的概念, 而流形的具有不变性的定义则在后面才引进.

在向量空间理论的基础讨论中, 我们采取比习惯的处理方法更加谨慎和精确的态度. 实际上, 这种语言的精确水准在后面的章节中并没有保持下去. 我们感觉到在概念如此清晰, 公理又如此熟悉的线性代数里, 从教学的角度看, 它对讲解各种微妙之处是精湛而彻底的, 譬如区分按常规为等同的空间, 讨论各种映射的本质属性, 等等. 自此之后, 当过度精确的语言会使得叙述更加繁琐时, 读者应该能够在发现某个论断写得不太严谨的情形下自己做出较为准确的叙述. 类似地, 在前几章中给出的证明是相当正式和细致的. 我们的原则仍然是, 一旦学生们掌握了那些构成了正式数学证明的观念, 那么使用通常的数学口语系统来给出论证则是安全和更为方便的了.

在形式化程度减少的同时, 数学的成熟程度并没有降低. 因此引进了具有不断增长的抽象性和高级的数学对象. 从我们的经验来说, 第九章包含了学生们最难于吸收的一些概念, 特别是流形的切空间的观念以及关于向量场的各种对象的李导数的观念.

本书从头到尾都安排了许多不同类型的习题. 其中一些是那种具有常规应用的性质. 另外一些则要求读者进行补充或推广正文中给出的结果的各种证明. 有时整个专题, 譬如傅里叶变换, 或留数计算, 以习题的形式出现. 由于教科书内容的十分抽象的特性, 我们强烈地劝告同学们要做出尽可能多的习题.

任何一件这类性质的事都极大地归功于除作者外的许多人, 但我们特别希

望感谢 L. Ahlfors, A.Gleason, R. Kulkami, R. Rasala 和 G. Mackey 的帮助, 以及 Dieudonné的书的总的影响. 我们还要感谢 Jones & Bartlett 出版社工作人员对准备这次再版时的极其珍贵的帮助.

L. H. L.

S. S.

于马萨诸塞州剑桥

1968, 1989

目 录

第零章　导引

在这一预备章里, 包含了构成当今数学思想基础的集合论的一个简单阐述. 我们由逻辑的简单讨论开始, 从而可以较精确地讨论集合论, 继而对数学对象可以被定义为集合的这种方式给予评论。在本章结尾的四节中, 我们处理了特殊的集合论专题.

我们的意图为将这个材料主要用作参考, 其中一些是读者熟悉的, 而另一些则可能是新的. 我们建议读者在读本章时, 首先 "快速地" 浏览一下, 然后再根据需要重读各细节.

0.1　逻辑: 量词

一个*语句*为一个句子, 根据它自身确定是真的或假的. 因此 '1<2' 与 '4+3=5' 分别为真的与假的数学语句. 数学中有许多句子含有变量, 所以其自身不代表真或假, 但当给予变量数值变量后, 就变成语句了. 简单的例子为 '$x < 4$', '$x < y$', 'x 为一个整数','$3x^2 + y^2 = 10$'. 这种句子被称为*语句框架*.若 $P(x)$ 为包含一个变量 'x' 的框架, 则 $P(5)$ 为在 $P(x)$ 中将 'x' 换成数字 '5' 后得到的语句. 例如, 若 $P(x)$ 为 '$x < 4$', 则 $P(5)$ 就是 '$5 < 4$', $P(\sqrt{2})$ 为 '$\sqrt{2} < 4$' 等等.

由框架 $P(x)$ 得到语句的另一途径为断言 $P(x)$ 总是真的. 我们用预先固定的句子 '对于每一个 x' 的方法来这样做. 所以 '对于每一个 x, $x < 4$' 是一个假

语句, 而 '对于每一个 x, $x^2-1=(x-1)(x+1)$' 是一个真语句. 这个预先固定的句子称为一个*全称量词*.同义语的句子为 '对每一个 x' 与 '对所有 x'. 它们的习惯使用记号为 '$(\forall x)$'. 我们可以用这种方法中的任意一种来读它. 我们经常认为不明确写出全称量词时, 包含变量的句子总是真的. 例如对于数的加法结合律经常被写作

$$x+(y+z)=(x+y)+z.$$

此处我们了解为对于所有 x,y 与 z, 方程均为真的语句. 因此实际的句子应该是

$$(\forall x)(\forall y)(\forall z)[x+(y+z)=(x+y)+z].$$

最后, 我们可以将框架 $P(x)$ 转化为一个语句, 用断言它有时是真的来实现. 我们将它记作 '存在一个 x 使 $P(x)$'. 这个过程称为*存在量词*. 这里同义的预先固定句子为 '有一个 x 使','对于某 x'. 我们用记号 '$(\exists\, x)$' 来表示.

语句 '$(\forall\, x)(x<4)$' 中仍包含变量 'x', 当然, 'x' 不能*自由*取值. 我们现在称它为一个*约束变量*.粗略地说, 量化的变量是约束的而非量化的变量是自由的. 记号 '$P(x)$' 仅用于在讨论的句子中,'x' 是自由的时候.

现在假定我们有一个含有两个自由变量的句子 $P(x,y)$. 显然, 我们需要两个量词才能由这个句子得到一个语句. 这导致一个非常重要的思考.*若两种类型的量词均被用到, 则它们的书写次序将影响语句的含义*; $(\exists\, y)(\forall\, x)P(x,y)$与$(\forall\, x)(\exists\, y)P(x,y)$*说的是不同的事情*. 第一种情况是说可以找到一个 y 使对所有 x 皆 :"存在一个 y 使对所有 $x\cdots$". 第二种情况则是说对于每一个 x, 可以找到一个 y 使:"对于每一个 x, 存在 y 使 $\cdots$". 但在第二种情况下, 非常可能发生这样的事情, 即当 x 改变时, 可以找到的 y 亦改变了. 所以对于所有 x, 一个单独 y 的存在性则为更强的语句. 例如, $(\forall\, x)(\exists\, y)(x<y)$ 是真的, 而 $(\exists\, y)(\forall\, x)(x<y)$ 是假的. 读者必须绝对弄清楚这一点; 这将关系到他的整个数学生涯. 第二个语句说, 存在一个 y , 记为 y_0, 使 $(\forall\, x)(x<y_0)$, 即使每一个数皆小于 y_0. 这是假的; 特别 y_0+1 就不小于 y_0. 第一个语句说, 对于每一个 x, 我们可以找到一个*对应的*y. 因此我们可以取 $y=x+1$.

另一方面, 同样类型的一组量词中, 次序并不影响其含义. 所以 '$(\forall\, x)(\forall\, y)$' 与 '$(\forall\, y)(\forall\, x)$' 具有同样的含义. 我们常常仅用一次量词记号来简化这种类似的量词丛, 例如 '$(\forall\, x,y)$', 它可以读作 '对于每一个 x 与 y'. 因此严格正确的 '$(\forall\, x)(\forall\, y)(\forall\, z)[x+(y+z)=(x+y)+z]$' 就有了较为习惯的写法 '$(\forall\, x,y,z)[x+(y+z)=(x+y)+z]$'. 对于存在量词组, 显然亦有同样的情况.

初学的学生一般会感到预先固定的句子 '对于每一个 x, 皆存在一个 y 使' 与 '存在一个 y 使对于每一个 x' 似乎是人为的且不习惯. 的确是这样的, 但是这种麻烦是为了使量词的次序得到固定而付出的代价, 从而使量词语句的含义是清楚与不含糊的. 量词在通常惯用的记述中是出现的, 但它们在惯用出现中, 常常蕴涵了模糊. 下面的两个句子就是这种含糊用法的好例子: "每一个 x 小于某个 y" 与 "某一个 y 大于每一个 x" . 若有一个测验所, 将会发现街上大多数人会认为这两个句子说的是一件事, 但是一半人将会感到这个共同的断言是假的, 而另一半人则认为它是真的! 这里的混乱是由于一个量词接着另一个量词的模型而产生的. 初学的读者不知道哪个取作里面的, 或第一个使用的量词. 我们第一个句子的两种可能的记号写法为 '$[(\forall\ x)(x < y)](\exists\ y)$' 与 '$(\forall\ x)[(x < y)(\exists\ y)]$', 它们分别为假的与真的. 数学家确实在更为习惯的写作中, 使用疑悬量词, 但这仅当它们已确定, 或者从上下文关系, 或者在与标准用法相比较中, 读者已理解了它们的应用次序. 一般说来, 一个疑悬量词将读作里面的或第一个应用的量词, 及由此来了解两个含糊的句子在哪个次序下变成真的与假的.

作过这个解释之后, 读者应该可以认可序列收敛的定义, 它涉及到三个量词且表述如下: 若 $(\forall\varepsilon)(\exists N)(\forall n)$(当 $n > N$, 则 $|x_n - x| < \varepsilon$), 则序列 $\{x_n\}$ 收敛于 x. 用完全同样的方式, 我们定义一个函数 f 在一点 a 是连续的, 这只要 $(\forall\varepsilon)(\exists\delta)(\forall x)$(当 $|x-a| < \delta$ 时, 则 $|f(x) - f(a)| < \varepsilon$). 我们经常用展示最后的框架来略去一个里面的全称量词, 使全称量词能够理解. 因此我们这样定义 f 在 a 是连续的: 若对于每一个 ε, 皆存在一个 δ 满足

$$\text{当}|x-a| < \delta, \text{则}|f(x) - f(a)| < \varepsilon.$$

我们将在以后学习这些定义. 我们仅注意到, 构筑对这些量词语句及类似量词语句的真正含义的一个直观理解是完全可能的.

0.2 逻辑连接词

当字 '与' 夹在两个句子之中时, 则当这两个组成的句子都是真的, 这个句子才是真的, 否则就是假的. 这就是说, 复合句子的 "真值"T 或 F 仅依赖于组成句子的真值. 因此我们可以用简单的 "真值表" 来描述将 '与' 字作用到组成句子的

结果:

P	Q	P 与 Q
T	T	T
T	F	F
F	T	F
F	F	F

此处 'P' 与 'Q' 表示任意语句框架. 如同 '与' 这样的字称为逻辑连接词. 用记号来表示连接词是很方便的. '与' 的一个标准记号为 '&'. 因此 'P & Q' 就读作 'P 与 Q'.

另一个逻辑连接词为字 '或'. 不幸的是, 在通常的论述中, 这个字的使用是含糊的, 它有时用作不可兼的意思, 其中 'P 或 Q' 表示 P 与 Q 之一是真的, 但不都是真的. 而有时, 它用来表示可兼的意思, 即至少有一个是真的, 又可能两个都是真的. 数学不允许本质上的含糊. 在数学中, '或'总是用于后一种情况, 因此我们有真值表

P	Q	P 或 Q
T	T	T
T	F	T
F	T	T
F	F	F

上面两个连接词都是二元的, 其含义为它们将两个句子组成一个新的句子. 字 '非' 作用于一个句子, 从而它实在不应该当作一个连接词; 然而, 它称为一个一元的连接词. 关于 '非' 的一个标准记号为 '$\sim$', 它的真值表显然为

P	$\sim P$
T	F
F	T

在习惯使用时, 字 '非' 一般是隐蔽于一个句子之内. 我们记作 'x 不等于 y', 而不宁愿写成 '非 (x 等于 y)'. 但是为了逻辑操作的目的, 否定记号 (字 '非' 或一个如同 '$\sim$' 的记号) 表示这个句子已经被否定. 当然, 我们将继续去使用 '$x \neq y$', 但必须铭记于心的是, 这是 '非 $(x = y)$' 或 '$\sim (x = y)$' 的习惯用法.

我们现在转入麻烦的连接词 '若 $\cdots$, 则 $\cdots$', 我们将写作 '若 P, 则 Q' 或 '$P \Rightarrow Q$'. 这几乎总是在全称量词 $(\forall\, x)(P(x) \Rightarrow Q(x))$ 中加以使用, 它的含义最

好用这种应用的案例来阐明. 我们考虑 '若 $x<3 \Rightarrow x<5$' 为一个真的句子. 更准确地说, 这对于所有 x 皆成立, 所以全称量词 $(\forall\ x)(x<3 \Rightarrow x<5)$ 是一个真的语句. 特别, 这一结论迫使我们同意 '$2<3 \Rightarrow 2<5$','$4<3 \Rightarrow 4<5$', 及 '$6<3 \Rightarrow 6<5$' 都是真的语句. 因此关于 '$\Rightarrow$' 的真值表包含下面的值.

P	Q	$P \Rightarrow Q$
T	T	T
T	F	—
F	T	T
F	F	T

另一方面, 我们考虑 '$x<7 \Rightarrow x<5$' 为一个假的句子, 所以我们就同意 '$6<7 \Rightarrow 6<5$' 是假的, 则对于 $P \Rightarrow Q$, 上面表的最后一列应有值 'F'.

如同我们已经考虑过的, 框架变量与逻辑连接词的综合称为*真值函数形式*.我们进一步用连接词综合诸如 '$P \Rightarrow Q$' 与 '$\sim P$' 这种初等形式来构造诸如 '$\sim(P \Rightarrow Q)$' 与 '$(P \Rightarrow Q)\ \&\ (Q \Rightarrow P)$' 的合成形式. 若一个句子可以由替换形来得到, 则这个句子就有一个给定的 (真值函数) 形式. 因此,'$x<y$ 或 $\sim(x<y)$' 就有形式 'P 或 $\sim P$', 这是由于它可以由这个形式, 将句子 '$x<y$' 替换句子变量 'P' 而得到. 复合真值函数形式有真值表, 它可以由综合初等表而得到. 例如, '$\sim(P \Rightarrow Q)$' 有下表, 整个形式的真值位于最后用到的连接词 (这个例子中为 $\sim$) 下面的列中.

P	Q	$\sim(P \Rightarrow Q)$	
T	T	F	T
T	F	T	F
F	T	F	T
F	F	F	T

因此仅当 P 为真及 Q 为假时, $\sim(P \Rightarrow Q)$ 才是真的.

若一个真值函数形式, 如 'P 或 $(\sim P)$', 总是真的 (即在真值表的最后一列中仅有 'T'), 则称为一个*重言式*或一个*重言形式*.读者可以验证

$$(P\&(P \Rightarrow Q)) \Rightarrow Q \quad 与 \quad ((P \Rightarrow Q)\&(Q \Rightarrow R)) \Rightarrow (P \Rightarrow R)$$

都是重言形式. 事实上, 任何不涉及量词的论证的正确原则必定被一个重言形式表示出来.

‘当且仅当’形式‘$P \Leftrightarrow Q$’, 或‘P 当且仅当 Q’, 或‘P iff Q’ 是‘$(P \Rightarrow Q)\&(Q \Rightarrow P)$’ 的缩写. 它的真值表为

P	Q	$P \Leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

即当 P 与 Q 有同样的真值时, $P \Leftrightarrow Q$ 才是真的, 否则就是假的.

若两个真值函数形式 A 与 B 在它们的直值表 (的最后一列) 相同, 则称 A 与 B 是等价的. 鉴于 ‘$\Leftrightarrow$’ 的表, 我们得知若 $A \Leftrightarrow B$ 是重言的, 则 A 与 B 等价, 其逆亦真. 将一个由形式 A 经替换得到的句子换成一个由同样替换等价形式 B 得来的句子是一个逻辑推理中常用的手段. 由于 ‘P’ 与 ‘$\sim(\sim P)$’ 为等价形式, 所以要证明一个语句 P 是真的, 只要证明 $\sim P$ 是假的即可. 其他的等价性为:

$$\sim(P \text{ 或 } Q) \Leftrightarrow (\sim P)\&(\sim Q),$$
$$(P \Rightarrow Q) \Leftrightarrow Q \text{ 或 } (\sim P),$$
$$\sim(P \Rightarrow Q) \Leftrightarrow P\&(\sim Q).$$

有一点习惯用法, 即我们满足于方便地在定义中用 ‘若’ 来代替正确的 ‘当且仅当’. 我们定义 f 在 x 连续, 假定如此这般. 当然, 含义就是 f 在 x 连续,当且仅当, 如此这般. 这不会引起困难, 这里由于显然 ‘当且仅当’ 表示一个定义已经给出了.

0.3　量词的否定

‘$\sim(\forall x)$’ 与 ‘$(\exists x)\sim$’ 具有同样的含义. 这种等价性可以用来移动一个否定记号过一串量词中的每一个量词, 这给出了下面重要的应用法则:

要取一个由词系构成的语句之否定, 我们只要简单地将每一个量词换成相反的量词, 并将否定词移至量词系之末即可.

因此

$$\sim(\forall x)(\exists y)(\forall z)P(x,y,z) \Leftrightarrow (\exists x)(\forall y)(\exists z)\sim P(x,y,z).$$

还有其他量词的推导原则可以被分离出来, 我们将在适当的时候提出来, 无需在此加以罗列.

0.4 集合

定义每一个数学对象为某种事物的一个集合是当今的常用方法，所以无论怎样简略，我们都必须检验这个基本概念.

一个集合为一些事物的全体，它的自身亦被看作是一个实物. 集合中的事物称为集合的*元素*或*成员*.'一个成员属于' 的记号为 '$\in$'(这是大写 ε 的样子)，所以 '$x \in A$' 读作 "x 是 A 的一个成员"，"x 属于 A" 或 "x 在 A 中".

在数学中，我们用相等记号 '=' 来表示逻辑恒等; $A = B$ 表示 A 就是 B. 现在一个集合 A 被当成与一个集合 B 为同样的事物当且仅当 A 与 B 恰好有相同的成员，即 '$A = B$' 表示 $(\forall\, x)(x \in A \Leftrightarrow x \in B)$.

若集合 A 的每一个元素都是 B 的一个元素，则称 A 为 B 的一个*子集合*或 A*包含于* B 中 (或 B 是 A 的一个*超集*). 包含的记号是 '$\subset$'. 因此 '$A \subset B$' 表示 $(\forall\, x)(x \in A \Rightarrow x \in B)$. 显然有

$$(A = B) \Leftrightarrow (A \subset B)\ \&\ (B \subset A).$$

这是常用的建立集合恒等式的一个方法: 我们用证明 $A \subset B$ 及 $B \subset A$ 来证明 $A = B$. 如果读者思考上面的等价性，他将看到它首先依赖于真实函数形式 '$P \Leftrightarrow Q$' 与 '$(P \Rightarrow Q)\ \&\ (Q \Rightarrow P)$' 的等价性，然后是 '$(\forall\, x)(R\ \&\ S)$' 与 '$(\forall\, x)R\ \&\ (\forall x)S$' 之间的明显量词等价性.

我们用说明其成员的方法来定义一个集合. 若集合是有限的，则其成员可以真正地被列出来，我们用的记号为将列出的成员用一个括号括起来. 例如 $\{1, 4, 7\}$为包含 3 个数 1, 4, 7 的集合，$\{x\}$为 x 的*一元集合*(仅含一个事物 x 为成员的集合) 及 $\{x, y\}$ 为 x 与 y 的*二元集合*. 我们可以不恰当地借用这个记号来表示某些无穷集合. 因此 $\{2, 4, 6, 8, \cdots\}$确实可以被当作所有偶正整数的集合. 但无穷集合一般都用语句框架来定义. 若 $P(x)$ 为一个包含自由变量 'x' 的框架，则 $\{x : P(x)\}$ 就是使 $P(x)$ 为真的所有 x 的集合. 换言之，$\{x : P(x)\}$ 为满足

$$y \in A \Leftrightarrow P(y)$$

的集合 A. 例如，$\{x : x^2 < 9\}$ 为所有适合 $x^2 < 9$ 的实数 x 的集合，即开区间 $(-3, 3)$，及 $y \in \{x : x^2 < 9\} \Leftrightarrow y^2 < 9$. 一个语句框架可以设想为叙述一个性质，一个事物可以具备或可以不具备它，从而 $\{x : P(x)\}$ 为所有具备该性质的事物之集合.

我们需要空集$\varnothing$, 它很相似于在算术中我们需要零一样. 若 $P(x)$不会是真的, 则 $\{x : P(x)\} = \varnothing$. 例如

$$\{x : x \neq x\} = \varnothing.$$

当我们在早些时候说, 所有的数学对象都被习惯地当作集合时, 这里认为读者已了解一个事物与该事物的名称之间的区别. 为了稳妥起见, 我再说几句, 一把椅子与字 '椅子' 不是同样的东西; 数 4 为一个数学对象, 它与数字 '4' 不是同样的东西. 数字 '4' 是数 4 的名字, 它与 '四','2+2' 及 'IV' 是一样的. 由现在观点看, 4 本身可以被当作某特定集合. 在本教程里, 我们没有必要将逻辑分析引向深入, 但有些读者可能有兴趣知道, 我们通常定义 4 为 {0,1,2,3}. 类似地, $2 = \{0, 1\}$, $1 = \{0\}$, 及 0 为空集 $\varnothing$.

由上述讨论及我们作的说明应该是清楚的, 即我们一个引号围住一个记号来表明记号之名称 (记号本身为其他某个东西的名称). 所以 "4" 为 '4' 的名称 (它本身是数 4 的名称). 这是严格正确的用法. 但数学家几乎常常错用了它. 精确的书写为: 命 x 为一个数; 称这个数为 'x'. 但这后者几乎总是被写成: 称这个数为 x. 这个不精确性对于阅读的数学家来说并不会引起困难, 而且它还消除了像雨一样的引号, 使印刷页得到节省. 但是这种记号的含糊处理使人成为可能的受害者. 这个人没有了解数学不关注于记号, 而是记号所表示的事物. 由于现在的读者已经安全地避开了这个陷阱, 所以我们可以放松一些, 并偶尔略去严格必要的引号.

为了避免过多使用 '集合' 这个字, 我们用许多同义字, 诸如 '类','全体','系', 与 '全部'. 因此我们可以说, "命 $\mathfrak{a}$ 为一个集合的类的系". 若一个鞋店是鞋的全体, 则一个鞋店连锁店就是这样一个 3 重事物.

0.5 限制变量

数学中的变量通常不允许取所有事物作为值; 它仅可以取某集合的成员为值, 我们称它为变元的*定义域*.定义域有时是显式标出的, 但常常仅是隐式表示的. 例如字母 'n' 习惯地用来表示一个整数, 所以 '$(\forall\, n)P(n)$' 将自动地被读作 "对于每一个整数 n, $P(n)$". 但有时 n 只取作一个正整数. 在可能有的含糊与疑惑的情况下, 我们将明确地指出限制并记作 '$(\forall\, n \in \mathbb{Z})P(n)$', 此处 '$\mathbb{Z}$' 为所有整数所成集合的标准记号. 量词照字义读作 '对于所有 $\mathbb{Z}$ 中的 n', 或更为随意地,'对于每一个整数 n'. 类似地,'$(\exists\, n \in \mathbb{Z})P(n)$' 读作 "在 $\mathbb{Z}$ 中存在一个 n 使 $P(n)$" 或

"存在一个整数 n 使 $P(n)$". 注意记号 '$\in$' 在此读作前置词 '之中'. 上面的量词称为限制量词.

同法, 我们限定了集合形成的方式, 既可是隐式的也可是显式的, 如 '$\{n : P(n)\}$' 与 $\{n \in \mathbb{Z} : P(n)\}$, 两者都读作 "所有满足 $P(n)$ 之整数 n 的集合".

限制变量可以由

$$(\forall x \in A)P(x) \Leftrightarrow (\forall x)(x \in A \Rightarrow P(x)),$$
$$(\exists x \in A)P(x) \Leftrightarrow (\exists x)(x \in A \ \&\ P(x)).$$
$$\{x \in A : P(x)\} = \{x : x \in A \ \&\ P(x)\}.$$

尽管在包括明确的限制变量的句子中, 没有任何含糊之处, 但用上指标位置书写限制子句时, 它有时会帮助我们看清句子的结构. 例如, $(\forall \varepsilon^{>0})(\exists n^{\in \mathbb{Z}})$. 某些限制在 0.1 节中已经隐含了. 若读者同意 $(\forall\ x)(x^2 - 1 = (x-1)(x+1))$ 是真的, 他或许会取 x 为一个实数.

0.6 序对与关系

正如读者从解析几何中了解到的, 序对是基本工具. 按照我们的普遍原则, 序对取作某一集合, 但是在这里, 我们不再考虑它是个什么集合, 只要它满足关键的刻画特征的性质:

$$\prec x, y \succ = \prec a, b \succ \Leftrightarrow x = a \text{ 与 } y = b.$$

所以 $\prec 1, 3 \succ \neq \prec 3, 1 \succ$.

对应, 或关系, 或一个映射的特殊情况, 或函数的概念对于数学来说是基本的. 一个对应是一对事物, 使给定任意两个事物 x 与 y, 二元 (联) 组 $\prec x, y \succ$ 或者是或者不是对应. 一个特殊的对应 (关系) 一般由一个语句框架 $P(x, y)$ 来表示, 它有两个自由变量, 且 x 与 y 对应当且仅当 $P(x, y)$ 是真的. 给定任何关系 (对应), 则所有对应元素的序对 $\prec x, y \succ$ 所成的集合称为它的图.

一个关系是一个数学对象, 正如我们已经说过多次, 流行的用法是将每一个数学对象看作这种或那种类型的集合. 由于一个关系的图是一个 (序对的) 集合, 所以习惯上取图为关系即可.因此一个关系 (对应) 简单地就是一个序对集合. 若 R 是一个关系, 则我们说 x 与 y 有关系 R, 并记作 'xRy', 当且仅当 $\prec x, y \succ \in R$. 我们亦称在 R 之下, x 对应于 y. 一个关系 R 的序对之第一个元素构成的集合

称为 R 的定义域, 并记为 dom$R$ 或 $\mathfrak{D}(R)$. 因此

$$\mathrm{dom}R = \{x : (\exists y) \prec x, y \succ \in R\}.$$

第二个元素构成的集合称为 R 的值域:

$$\text{range } R = \{y : (\exists x) \prec x, y \succ \in R\}.$$

一个关系 R 的逆R^{-1} 为 R 中的序对的次序颠倒后得到的序对集合:

$$R^{-1} = \{\prec x, y \succ : \prec y, x \succ \in R\}.$$

一个具有二个自由变量的语句框架 $P(x,y)$ 实际上确定了一对相互为逆的关系 $R\&S$, 称为 P 的图, 如下所示:

$$R = \{\prec x, y \succ : P(x,y)\}, \quad S = \{\prec y, x \succ : P(x,y)\}.$$

一个二维变量框架及选定一个变量作为第一个可以称为一个有向框架. 于是一个有向框架对于它的图就有一个惟一确定的关系. 实数 $\mathbb{R}$ 上的严格不等式关系将被看作是集合 $\{\prec x, y \succ : x < y\}$, 这是由于 '$x < y$' 中的变量有一个自然次序.

所有序对的集合 $A \times B = \{\prec x, y \succ : x \in A \ \& \ y \in B\}$, 其中第一个元素属于 A 及第二个元素属于 B, 称为集合 A 与 B 的笛卡儿乘积. 一个关系 R 总是 dom$R$ × range$R$ 的一个子集. 若两个 "因子空间" 是同样的, 我们就可以用指数记号: $A^2 = A \times A$.

笛卡儿乘积 $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ 为 "解析平面".解析几何就是基于 $\mathbb{R}^2$ 与欧氏平面 $\mathbb{E}^2$(后者由一个坐标系决定) 之间的一一坐标对应, 它使得我们可以代数地来处理几何问题及几何地去处理代数问题. 特别地, 由于实数集合之间的一个关系是 $\mathbb{R}^2$ 的一个子集, 所以我们可以由对应的欧氏平面的子集, 或如本页所示任何欧氏平面的模型的子集将它 "画" 出来. 一个简单的笛卡儿乘积如图 0.1 所示 ($A \cup B$ 为集 A 与 B 的并).

若 R 是一个关系及 A 为任意集合, 则 R 关于 A 的限制, $R \restriction A$, 为 R 中第一个元素属于 A 的序对构成之子集:

$$R \restriction A = \{\prec x, y \succ : \prec x, y \succ \in R \ \& \ x \in A\}.$$

因此 $R \restriction A = R \cap (A \times \text{range}R)$, 此处 $C \cap D$ 为集合 C 与 D 的交.

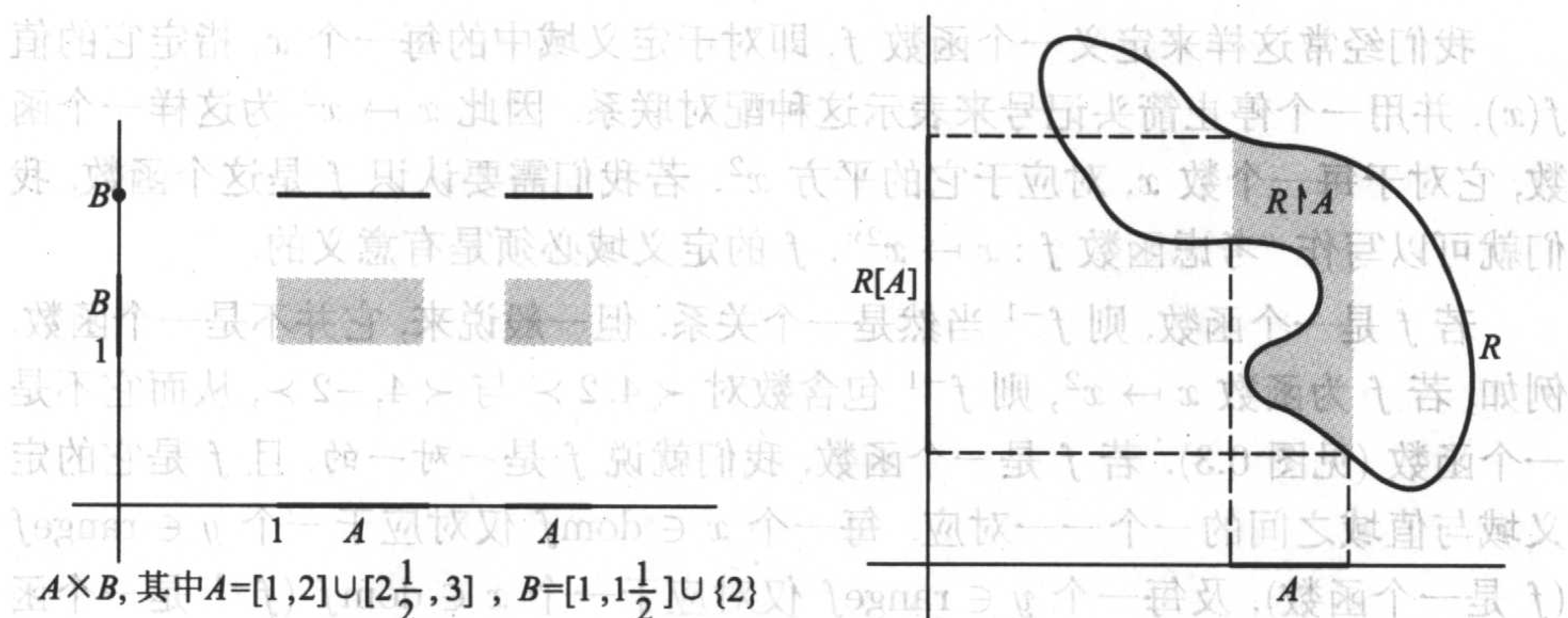


图 0.1　　图 0.2

若 R 是一个关系及 A 为任意集合, 则在 R 之下 A 的像,$R[A]$, 为 R 中序对的第二个元素构成的集合, 其中第一个元素属于 A:

$$R[A] = \{y : (\exists x)(x \in A \ \& \prec x, y \succ \in R)\}.$$

因此 $R[A] = \mathrm{range}(R|A)$, 如图 0.2 所示.

0.7　函数与映射

一个*函数*是一个关系 f, 它使定义域中每一个元素 x 正好对应于一个值域元素 y. 这个性质可以表示如下:

$$\prec x, y \succ \in f \quad \text{及} \quad \prec x, z \succ \in f \Rightarrow y = z.$$

如此由 f 与 x 确定的 y 记为 $f(x)$:

$$y = f(x) \Leftrightarrow \prec x, y \succ \in f.$$

人们常常会设想一个函数是积极的而一个非函数的关系是被动的. 一个函数作用于它的定义域中一个元素 x 即给出 $f(x)$. 我们取 x 并将 f *作用*于它; 事实上, 我们常常称一个函数为一个*算子*. 另一方面, 若 R 是一个关系, 但不是一个函数, 则相对于它的定义域中的一个元素 x, 一般并不存在与之对应的特别的 y, x 与 y 的配对被视为比较被动的.

我们经常这样来定义一个函数 f, 即对于定义域中的每一个 x, 指定它的值 $f(x)$, 并用一个停止箭头记号来表示这种配对联系. 因此 $x \mapsto x^2$ 为这样一个函数, 它对于每一个数 x, 对应于它的平方 x^2. 若我们需要认识 f 是这个函数, 我们就可以写作 "考虑函数 $f: x \mapsto x^2$". f 的定义域必须是有意义的.

若 f 是一个函数, 则 f^{-1} 当然是一个关系. 但一般说来, 它并不是一个函数. 例如, 若 f 为函数 $x \mapsto x^2$, 则 f^{-1} 包含数对 $\prec 4,2 \succ$ 与 $\prec 4,-2 \succ$, 从而它不是一个函数 (见图 0.3). 若 f 是一个函数, 我们就说 f 是一对一的, 且 f 是它的定义域与值域之间的一个一一对应. 每一个 $x \in \text{dom} f$ 仅对应于一个 $y \in \text{range} f$ (f 是一个函数), 及每一个 $y \in \text{range} f$ 仅对应于一个 $x \in \text{dom} f$ (f^{-1} 是一个函数).

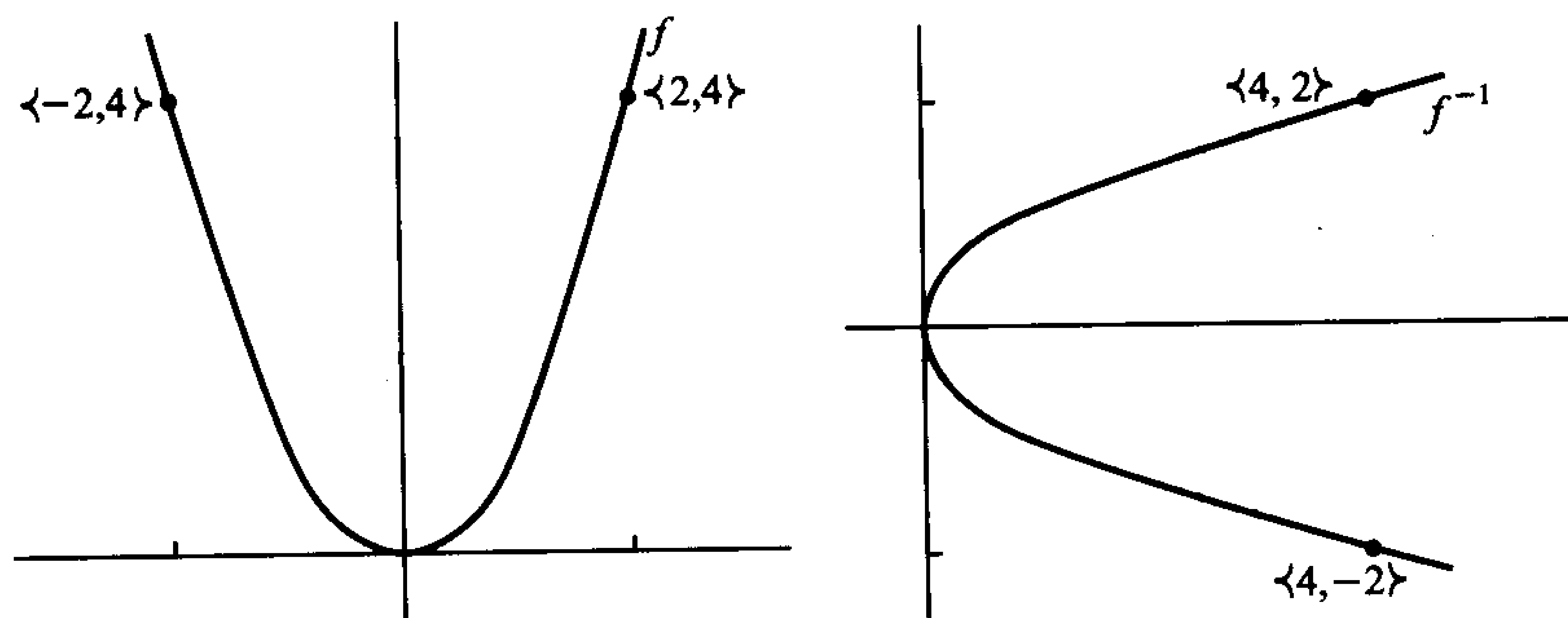


图 0.3

记号

$$f: A \to B$$

读作 "一个 A 上至 B 的函数", 或 "由 A 到 B 的函数". 这个记号表示 f 是一个函数, 其 $\text{dom} f = A$ 及 $\text{range} f \subset B$. 许多人感到函数的真正概念应该包括所有这些成分; 即一个函数应当被当成一个有序的三元组 $\prec f, A, B \succ$, 其中 f 为根据我们有限制定义的一个函数, A 为 f 的定义域, 及 B 为 f 值域的一个超集, 在本书中, 我们称它为 f 的值域. 我们将用术语 '映射','映入', 及 '变换' 来表示这样的一个三元组, 从而记号 $f: A \to B$ 从整体上表示一个映射. 进而言之, 当什么集合是值域没有疑问时, 我们常常称函数 f 本身为一个映射, 这是由于三元组 $\prec f, A, B \succ$ 是被 f 决定的. 两个箭头记号可以联合为, 如 "由 $x \mapsto x^2$ 定义 $f: \mathbb{R} \to \mathbb{R}$".

若 f 是一对一的, 则映射 $f: A \to B$ 就称为单射的.若 $\text{range}f = B$, 则称为满射的. 若它既是单射又是满射的, 则称为双射的. 因此一个双射的映射 $f: A \to B$ 是它的定义域 A 与值域 B 之间的一个一一对应. 当然一个函数总是到它值域 R 的一个满射, 从而语句 f 是满射的意思就是 $R = B$, 其中 B 理解为值域.

0.8 积集; 指标记号

近代数学的特征习性之一是当一种新的事物被定义及稍作讨论后, 人们就会立刻看看所有这种事物的集合. 随着由 A 至 S 的一个函数的概念很好地被掌握后, 我们自然地会考虑所有由 A 至 S 的函数集合. 我们将这个集合记为 S^A. 因此 $\mathbb{R}^\mathbb{R}$ 就是所有一个实变量的实函数所成的集合, 及 $S^{\mathbb{Z}^+}$ 为 S 上所有无穷序列所成之集合 (我们理解一个无穷序列就是一个函数, 其定义域为所有正整数所成之集合 $\mathbb{Z}^+$). 类似地, 若我们置 $\overline{n} = \{1, \cdots, n\}$, 则 $S^{\overline{n}}$ 为 S 上所有长度为 n 的有限序列所成之集合.

若 B 是 S 的一个子集, 则 (相对于 S) 它在 S 上的特征函数, 一般地记为 χ_B, 在 B 上取常数值 1, 而在 B 以外则取常数值 0. 所以, S 的子集的所有特征函数的集合为 2^S(由于 $2 = \{0, 1\}$). 但是由于这个函数集合与 S 的子集所成的集合是自然地一一对应的, χ_B 对应于 B, 所以我们将这两个集体恒等起来, 因此 2^S 亦被解释为 S 所有子集所成之集合. 我们将用本节所剩下的大部分篇幅来讨论数学允许的类似定义的含糊性.

有序三元组 $\prec x, y, z \succ$ 通常被定义为有序数对 $\prec\prec x, y \succ, z \succ$. 这个定义的理由可能是二个变量 x 与 y 的一个函数通常被看成是单个数对 $\prec x, y \succ$ 的一个函数, 所以, 例如, 一个两个实变量的实函数是 $(\mathbb{R} \times \mathbb{R}) \times \mathbb{R}$ 的一个子集. 但我们亦将这样一个函数当成笛卡儿三维空间 $\mathbb{R}^3$ 的一个子集. 因此, 我们定义 $\mathbb{R}^3$ 为 $(\mathbb{R} \times \mathbb{R}) \times \mathbb{R}$, 即我们定义: 有序三元组 $\prec x, y, z \succ$ 为 $\prec\prec x, y \succ, z \succ$.

另一方面, 有序三元组 $\prec x, y, z \succ$ 亦可以被当成有限序列 $\{\prec 1, x \succ, \prec 2, y \succ, \prec 3, z \succ\}$, 它当然是一个不同的事物. 一个有序三元组的这两种模型都同样好地起着作用, 而数学家又忽略了它们的区别. 当我们在以后讨论自然同构时 (1.6 节), 我们将更多地谈到这一点. 暂时, 我们简单地将 $\mathbb{R}^3$ 与 $\mathbb{R}^{\overline{3}}$ 当成是同样的; 一个有序三元组可以或者被 "看作" 一个序对, 其中第一个元素亦为一个序对, 或者为一个长度 3 的序列 (或者, 作为一个序对, 其中第二个元素为一个序对).

类似地, 我们设想笛卡儿四维空间 $\mathbb{R}^4$ 为 $\mathbb{R}^{\bar{4}}, \mathbb{R}^2 \times \mathbb{R}^2$, 或 $\mathbb{R}^1 \times \mathbb{R}^3 = \mathbb{R}^1 \times ((\mathbb{R} \times \mathbb{R}) \times \mathbb{R})$, 等等. 很明显, 我们实际上假定了关于序对的一个结合律, 但我们并不真正具有它.

这种含糊性, 即我们欲将两个真正相异的事物等同起来, 是确切决定事物是什么东西的一个必要推论. 这是我们为集合论的精确性所付的代价之一; 当数学不明确的时候, 就会只有一个简单的模糊概念.

数学中常用的指标设计亦会有含糊的关系, 我们应该加以检查. 一个有指标的集体, 作为一个集合, 不是别的, 它是一个函数, 即指标函数的值集合. 一个特殊的有指标的事物 x_i, 只是简单地表示函数在定义域中元素 i 的函数值. 如果指标集合为 I, 则有指标的集合就记为 $\{x_i : i \in I\}$ 或 $\{x_i\}_{i \in I}$(或 $\{x_i\}_{i=1}^{\infty}$, 其中 $I = \mathbb{Z}^+$). 总之, 这个记号建议我们将有指标的集合看作是将指标遍历指标集合 I, 然后将带有指标的事物搜集起来所得之集合, 即一个有指标的集合被看作是集合*加上* 指标函数. 这种双重性也反映在同样的记号常常表示映射. 所以我们提到*序列* $\{x_n\}_{n=1}^{\infty}$, 此处序列就是映射 $n \mapsto x_n$. 我们相信, 如果读者检查一下他关于序列的概念, 他将发现存在这个含糊性. 他认为既不正是集合, 亦不正是映射, 而是映射并强调它的值域, 或值域 "加上" 映射. 由于集合论不能用任何简单与优雅的方式来反映这个微小的差别, 所以我们将取一个有指标的集合为指标函数. 当然, 同样的值域事物可以被不同的指标加以重复; 一种指标是否一对一是没有关系的. 我们还要注意将指标加于已有指标的集合亦无限制; 每一个集合至少可以以自身为指标 (用恒等函数).

除了关于含糊的 '$\{x_i : i \in I\}$' 之外, 我们还没有通用的关于指标函数的记号. 由于 x_i 为函数在 i 的值, 所以我们可以设想 'x_i' 为 '$x(i)$' 的另一种写法, 在这种情况下, 我们将函数记为 'x' 或 '$\boldsymbol{x}$'. 当我们说,"考虑 n 元 (联) 组 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ$" 时, 我们确实在有序 n 元组的情况下用到 $\boldsymbol{x}$. 另一方面, 并没有必要的理由来使用这个记号. 我们可以称指标函数为任何我们想叫的东西; 若它是 f, 则当然对于所有的 i 有 $f(i) = x_i$.

我们现在回到笛卡儿乘积的一般定义. 较早时, 我们曾主张过 (在一个特殊情况下), 笛卡儿乘积 $A \times B \times C$ 为所有的有序三元组 $\prec x_1, x_2, x_3 \succ$ 的集合, 其中 $x_1 \in A, x_2 \in B$, 及 $x_3 \in C$. 一般地, $A_1 \times A_2 \times \cdots \times A_n$, 或 $\prod_{i=1}^n A_i$, 为有序 n 元组 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ$ 的集合, 此处 $x_i \in A_i, i = 1, \cdots, n$. 如果我们将一个有序 n 元组解释为 $\overline{n} = \{1, \cdots, n\}$ 上的一个函数, 则得

$$\prod_{i=1}^n A_i \text{为所有函数 } \boldsymbol{x} \text{ 的集合, 其中定义域 } \overline{n} \text{满足: 对于所有} i \in \overline{n}, x_i \in A_i.$$

这个定义几乎可以逐字地推广, 使我们得到任意一个有指标的集合集体之笛卡儿乘积概念.

定义 8.1 有指标的集合的集体 $\{S_i : i \in I\}$ 的笛卡儿乘积 $\prod_{i\in I} S_i$ 为所有这种函数 f 的集合, 其定义域为指标集 I, 并满足: 对于所有 $i \in I, f(i) \in S_i$.

关于乘积, 我们也可以用记号 $\prod\{S_i, i \in I\}$ 来表示, 并用 f_i 表示值 $f(i)$.

0.9 合成

如果我们给出映射 $f : A \to B$ 与 $g : B \to C$, 则 g 与 f 的合成 $g \circ f$ 是由 A 至 C 的映射, 其定义为

$$(g \circ f)(x) = g(f(x)), \quad \text{对于所有} x \in A \text{成立}.$$

这是初等微积分中一个函数算子的函数. 如果 f 与 g 为由 $f(x) = x^{1/3} + 1$ 与 $g(x) = x^2$ 定义的由 $\mathbb{R}$ 至 $\mathbb{R}$ 的映射, 则 $f \circ g(x) = (x^2)^{1/3} + 1 = x^{2/3} + 1$ 而 $g \circ f(x) = (x^{1/3} + 1)^2 = x^{2/3} + 2x^{1/3} + 1$. 注意为了使 $g \circ f$ 可以被定义, f 的值域必须是 g 的定义域. 这个算子可能是数学的基本二元算子.

引理 9.1 合成满足结合律:

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

证明 $(f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))) = (f \circ g)(h(x)) = ((f \circ g) \circ h)(x)$ 对于所有 $x \in \text{dom}h$ 成立. □

若 A 是一个集合, 则恒等映射 $I_A : A \to A$ 为将每一个 $x \in A$ 映至自身的映射. 因此 $I_A = \{\prec x, x \succ : x \in A\}$. 如 f 将 A 映射至 B, 则显然

$$f \circ I_A = f = I_B \circ f.$$

若 $g : B \to A$ 满足 $g \circ f = I_A$, 则我们称 g 为 f 的一个*左逆元*及 f 为 g 的一个*右逆元*.

引理 9.2 若映射 $f : A \to B$ 既有一个右逆元 h, 又有一个左逆元 g, 则它们必须相等.

证明 利用代数技巧与结合运算, 我们有

$$h = I_A \circ h = (g \circ f) \circ h = g \circ (f \circ h) = g \circ I_B = g. \qquad \square$$

在这种情况下, 我们称满足 $f \circ g = I_B$ 与 $g \circ f = I_A$ 惟一确定的映射 $g : B \to A$ 为 f 的*逆元*. 于是我们有

定理 9.1 *一个映射 $f : A \to B$ 有逆元当且仅当它是双射的, 在这种情况下, 它的逆元为它的关系逆元 f^{-1}.*

证明 如果 f 是双射的, 则关系逆元 f^{-1} 为一个由 B 至 A 的函数, 且方程 $f \circ f^{-1} = I_B$ 与 $f^{-1} \circ f = I_A$ 显然成立. 另一方面, 若 $f \circ g = I_B$, 则由于 B 中任何 y 都可以被写成 $y = f(g(y))$, 所以 f 是满射的. 又若 $g \circ f = I_A$, 则由方程 $f(x) = f(y)$ 可以导出 $x = g(f(x)) = g(f(y)) = y$, 所以 f 是单射的. 因此如果 f 有一个逆元, 则它是双射的. $\square$

现在命 $\mathcal{S}(A)$ 为所有双射映射 $f : A \to A$ 所成的集合. 则在合成二元运算之下, $\mathcal{S}(A)$ 是闭的, 而且

(1) $f \circ (g \circ h) = (f \circ g) \circ h$ 对于所有 $f, g, h, \in \mathcal{S}(A)$ 成立;

(2) 存在惟一的 $I \in \mathcal{S}(A)$ 使 $f \circ I = I \circ f = f$ 对所有 $f \in \mathcal{S}(A)$ 成立;

(3) 对于每一个 $f \in \mathcal{S}$, 皆存在惟一的 $g \in \mathcal{S}$ 满足 $f \circ g = g \circ f = I$.

在一个二元运算下具有这些性质的任何闭集合 G 称为关于这个运算的一个*群*. 因此 $\mathcal{S}(A)$ 是关于合成的一个群.

合成也可以用来定义关系如下: 若 $R \subset A \times B$ 及 $S \subset B \times C$, 则 $S \circ R \subset A \times C$ 的定义为

$$\prec x, z \succ \in S \circ R \Leftrightarrow (\exists y^{\in B})(\prec x, y \succ \in R \ \& \prec y, z \succ \in S).$$

若 R 与 S 为映射, 则这个定义与我们前面说的定义是一致的.

0.10 对偶性

存在另一种初等但重要的现象称为*对偶性*,实际上, 它在数学的所有分支中都出现. 命 $F : A \times B \to C$ 为任意两个变量的函数. 显然如果 x 固定, 则 $F(x, y)$ 是一个变量 y 的函数. 就是说, 对每一个固定的 x, 则有一个由 $h^x(y) = F(x, y)$ 定义的函数 $h^x : B \to C$. 所以 $x \mapsto h^x$ 是一个由 A 至 C^B 的映射 φ. 类似地, 每

一个 y 皆导致一个函数 $g_y \in C^A$, 此处 $g_y(x) = F(x,y)$ 及 $y \mapsto g_y$ 为一个由 B 至 C^A 的映射 θ.

假定现在反过来, 我们有一个映射 $\varphi: A \to C^B$. 对于每一个 $x \in A$, 按指标记号, 我们确定 φ 的对应值为 h^x, 所以 h^x 为由 B 至 C 的一个函数, 及我们由 $F(x,y) = h^x(y)$ 定义 $F: A \times B \to C$. 现在我们回到刚开始处. 因此映射 $\varphi: A \to C^B, F: A \times B \to C$, 及 $\theta: B \to C^A$ 是等价的, 而且可以设想为同样现象的三种不同途径. 最终的映射 φ 与 θ 称为互相对偶的.

映射 φ 是有指标的函数族 $\{h^x : x \in A\} \subset C^B$. 现在假定 $\mathfrak{F} \subset C^B$ 为一个由 B 至 C 的无指标函数集体, 并由 $F(f,y) = f(y)$ 定义 $F: \mathfrak{F} \times B \to C$. 则 $\theta: B \to C^{\mathfrak{F}}$ 由 $g_y(f) = f(y)$ 定义. 在这里发生的事只是在表达式 $f(y)$ 中, 我们把两个记号都当成变量了, 所以 $f(y)$ 是 $\mathfrak{F} \times B$ 上的一个函数. 因此当我们将 y 固定时, 我们就得到了 $\mathfrak{F}$ 上的一个将 $\mathfrak{F}$ 映射至 C 的函数.

作为主要进展, 我们将要看看这个对偶性原则的重要应用. 例如, 一个 $m \times n$ 矩阵为一个 $\mathbb{R}^{\overline{m} \times \overline{n}}$ 上的函数 $t = \{t_{ij}\}$. 我们将矩阵写成一个数的长方形列阵, 其中 'i' 为行指标及 'j' 为列指标, 从而 t_{ij} 为第 i 行与第 j 列相交处的那个数. 若我们将 i 固定, 则我们得构成 i 行的 n 元组, 于是矩阵可以被说成是一个行 n 元组的 m 元组. 类似地 (对偶性), 它也可以被看作列 m 元组的 n 元组.

同理, 一个由 A 至 B 的函数 n 元组 $\prec f_1, \cdots, f_n \succ$ 可以被当作由 A 至 $B^{\overline{n}}$ 的单个 n 元有值函数组

$$a \mapsto \prec f_1(a), \cdots, f_n(a) \succ .$$

在略为不同的应用中, 对偶性将让我们认为一个有限维的向量空间 V 是它自身的第二共轭空间 $(V^*)^*$.

由这个观点来观察初等几何是有益的. 今天我们把一条直线当成一个几何点集合. 一个较老与较为中立的观点则是将点与直线取作不同类型的本原物体. 所以, 命 A 为所有点所成的集合 (从而正如我们现在所看待的, A 就是欧氏平面), 及 B 为所有直线所成之集合. 命 F 为关联函数: 若 p 与 l 相关联 (p"在"l 上, l"在"p 上), 则 $F(p,l) = 1$, 否则 $F(p,l) = 0$. 所以 F 将 $A \times B$ 映至 $\{0,1\}$. 因此对于每一个 $l \in B$, 函数 $g_l(p) = F(p,l)$ 就是我们设想的直线 l 的点集合的特征函数(当 p 在 l 上时, $g_l(p) = 1$, 而当 p 不在 l 上时 , $g_l(p) = 0$). 从而每一条直线决定了一个位于它上面的点集合. 但是对偶地, 经过它的特征函数 $h^p(l)$, 每一个点决定了在它"上" 面的直线 l 的集合. 所以在完全对偶性中, 我们可以将直线当作一个点集合, 而将点看成一个直线之集合. 在射影几何中, 几何的这个对偶

性是基本的.

发明新的记号来表示“部分”函数有时是颇难的, 这种函数是由一个多变量函数中固定一个变量而得到的, 正如我们在上面所做的那样, 我们置 $g_y(x) = F(x,y)$. 但在这种情况下, 还有另一种常常是很有用的办法; 这是在“变动变量”处放一个点. 因此 $F(a,\cdot)$ 为一个变量的函数, 它是由 $F(x,y)$ 中, 将 x 固定为值 a 而得到的, 所以在对偶性的最初讨论中, 我们有

$$h^x = F(x,\cdot), \quad g_y = F(\cdot,y).$$

如果 f 是一个变量的函数, 则我们可以将它记为 $f(\cdot)$, 从而亦可以将上面的方程表示为 $h^x(\cdot) = F(x,\cdot), g_y(\cdot) = F(\cdot,y)$. 这个记号的缺点是我们不能指出未失去意义的代换. 因此函数 $F(x,\cdot)$ 在 b 点之值为 $F(x,b)$, 但从这个计算, 我们不能够反向来读, 并告诉我们, 什么函数是被计算的, 从而迫使我们用某些麻烦的记号, 例如 $F(x,\cdot)|_b$, 这颇难驾御. 然而在没有计算困难时, 点记号的使用仍然是有好处的. 除了消除如上所说的临时记号的需要外, 它在下面情况下仍然可以被使用, 即让我们立刻看到变量的位置, 但严格地讲, 这是多余的.

例如, $D_\xi F$ 在以后将表示在 (固定的) 方向 ξ 上的方向导数. 这是一个函数, 它在 α 处的值为 $D_\xi F(\alpha)$, 而记号 $D_\xi F(\cdot)$ 使这个隐式理解的事实明显了.

0.11 布尔运算

命 S 为一个固定的定义域, 及 $\mathfrak{F}$ 为 S 的一个子集合系. 至少属于 $\mathfrak{F}$ 中一个集合的所有元素所成的集合称为 $\mathfrak{F}$ 的并, 或 $\mathfrak{F}$ 中所有集合的并. 我们将它记为 $\cup\mathfrak{F}$ 或 $\cup_{A\in\mathfrak{F}}A$. 因此我们有

$$\cup\mathfrak{F} = \{x : (\exists A^{\in F})(x \in A)\}, \quad y \in \cup\mathfrak{F} \Leftrightarrow (\exists A^{\in\mathfrak{F}})(y \in A).$$

我们常常考虑系 $\mathfrak{F}$ 已经有指标了, 即我们假定给出一个集合 I(指标的集合), 及一个由 I 至 $\mathfrak{F}$ 的满射映射 $i \mapsto A_i$, 所以 $\mathfrak{F} = \{A_i : i \in I\}$. 从而有指标集合的并记为 $\cup_{i\in I}A_i$ 或 $\cup\{A_i : i \in I\}$. 指标的设计有技术上与心理上的优势, 我们将经常使用它.

如果 $\mathfrak{F}$ 是有限的, 并假定它本身或者指标集已被列举, 则对于它的并, 还有一个不同的记号可用. 若 $\mathfrak{F} = \{A,B\}$, 则我们记并为 $A \cup B$, 这是一个用列举名字来表示的记号. 注意在此我们有 $x \in A \cup B \Leftrightarrow x \in A$ 或 $x \in B$. 若

$\mathfrak{F}=\{A_i : i = 1, \cdots, n\}$, 则对于 $\cup\mathfrak{F}$, 我们一般将它写为 '$A_1 \cup A_2 \cup \cdots \cup A_n$' 或 '$\bigcup_{i=1}^n A_i$'.

有指标系 $\{A_i\}_{i\in I}$ 的交记为 $\cap_{i\in I}A_i$, 它是属于每一个 A_i 的点构成的集合. 所以

$$x \in \bigcap_{i\in I} A_i \Leftrightarrow (\forall_i^{\in I})(x \in A_i).$$

对于一个无指标的系 $\mathfrak{F}$, 我们用记号 $\cap\mathfrak{F}$ 或 $\cap_{A\in\mathfrak{F}}A$. 若 $\mathfrak{F} = \{A, B\}$, 则 $\cap\mathfrak{F} = A\cap B$.

S 的子集 A 的补集 (或余集)A' 为所有满足 $x \in S$, 但不属于 A 之元素所成的集合:$A' = \{x^{\in S} : x \notin A\}$. 摩尔根定律称一个交集的补集是补集之并集:

$$\left(\bigcap_{i\in I} A_i\right)' = \bigcup_{i\in I}(A_i').$$

这是 "否定量词规则" 的立即推论. 这里 '不总是在' 与 '有时不在' 之间的等价性: $[\sim (\forall\, i)(x \in A_i) \Leftrightarrow (\exists\, i)(x \notin A_i)]$, 确切地说为

$$x \in \left(\bigcap_{i} A_i\right)' \Leftrightarrow x \in \bigcup_{i}(A_i').$$

若我们置 $B_i = A_i'$, 并再取补集, 即得到对偶形式 $(\cup_{i\in I}B_i)' = \cap_{i\in I}(B_i')$.

其他量词原则可以推出规则如下: 由 $P \,\&\, (\exists\, x)Q(x) \Leftrightarrow (\exists\, x)(P \,\&\, Q(x))$ 可以推出

$$B \cap \left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I}(B \cap A_i),$$

还有

$$B \cup \left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I}(B \cup A_i),$$

$$B \cap \left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I}(B \cap A_i),$$

$$B \cup \left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I}(B \cup A_i).$$

对于两个集合的情况, 这些规律就是下面集合代数的熟知规律:

$$(A \cup B)' = A' \cap B', \quad (A \cap B)' = A' \cup B'(\text{摩尔根}),$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

即使在这里, 用指标的想法亦会使规律变得更为直观. 因此, 当想到作为 '不总是在' 与 '有时不在' 之间的等价性时,

$$(A_1 \cap A_2)' = A_1' \cup A_2'$$

就变得显然了.

若系 $\mathfrak{F}$ 中不同的集合都没有公共元素, 即如果 $(\forall\, X, Y^{\in \mathfrak{F}})(X \neq Y \Rightarrow X \cap Y = \varnothing)$, 则称系 $\mathfrak{F}$ 是*不相交的*.对于一个有指标的系 $\{A_i\}_{i \in I}$, 条件就变成 $i \neq j \Rightarrow A_i \cap A_j = \varnothing$. 若 $\mathfrak{F} = \{A, B\}$, 则我们就简单地说, A 与 B 是不相交的.

给予 $f : U \to V$ 及一个 V 的子集构成的有指标的系 $\{B_i\}$, 我们有下面重要的恒等式

$$f^{-1}\left[\bigcup_i B_i\right] = \bigcup_i f^{-1}[B_i], \quad f^{-1}\left[\bigcap_i B_i\right] = \bigcap_i f^{-1}[B_i],$$

及对于一个单个集合 $B \subset V$ 有

$$f^{-1}[B'] = (f^{-1}[B])'.$$

例如

$$x \in f^{-1}\left(\bigcap_i B_i\right) \Leftrightarrow f(x) \in \bigcap_i B_i \Leftrightarrow (\forall\, i)(f(x) \in B_i)$$

$$\Leftrightarrow (\forall i)(x \in f^{-1}[B_i]) \Leftrightarrow x \in \bigcap_i f^{-1}[B_i].$$

当 f 被换成任何 R 时, 上面所示的三个恒等式中的第一个, 而不是另外两个, 仍然是正确的. 由交换律可知 $(\exists\, x)(\exists\, y)A \Leftrightarrow (\exists\, y)(\exists\, x)A$. 由于 '$(\exists\, x)(\forall\, y)$' 与 '$(\forall\, y)(\exists\, x)$' 有不同含义, 所以第二个恒等式对一般 R 不成立.

0.12 分拆与等价关系

一个不相交的集合系 $\mathfrak{F}$, 其集合之并为 A, 则 $\mathfrak{F}$ 就称为 A 的一个*分拆*. 我们称 $\mathfrak{F}$ 的元素为 '纤维'并称 $\mathfrak{F}$*纤维了* A 或 $\mathfrak{F}$ 是 A 的一个*纤维表示*. 例如, 欧氏平面上平行于一条给定直线的直线集合就是平面的一个纤维表示. 如果 '$\bar{x}$' 表示含有

点 x 的惟一纤维, 则 $x \mapsto \overline{x}$ 是一个满射映射 $\pi: A \to \mathfrak{F}$, 我们称它为 A 在 $\mathfrak{F}$ 上的*投影*.由一个集合 A 至 A 的一个纤维 $\mathfrak{F}$ 是构成新的数学对象的主要途径之一.

任何函数 f 自动地将它的定义域纤维成一些集合, 在每个集合上, f 取常数值. 如果 A 是一个欧氏平面及 $f(p)$ 为在某坐标之下, 点 p 的 x 坐标, 则 f 在每一条垂直线上取常数值; 更确切地说, 对于 $\mathbb{R}$ 中每一个 x, $f^{(-1)}(x)$ 为一条垂直线. 进而言之, $x \mapsto f^{(-1)}(x)$ 为一个由 $\mathbb{R}$ 至所有纤维 (垂直线) 所成集合之间的一个双射映射. 一般说来, 若 $f: A \to B$ 为任何满射映射, 及若对于 B 中每一个值 y, 我们置

$$A_y = f^{(-1)}(y) = \{x \in A : f(x) = y\},$$

则 $\mathfrak{F} = \{A_y : y \in B\}$ 是 A 的一个纤维表示及 $\varphi : y \to A_y$ 是一个由 B 至 $\mathfrak{F}$ 的双射映射. 由于 $\varphi \circ f(x) = \varphi(f(x))$ 为 A 中所有适合 $f(z) = f(x)$ 的 z 的集合 $\overline{x}$, 所以 $\varphi \circ f$ 是一个投影 $\pi : A \to \mathfrak{F}$.

上面由 A 上一个函数生成 A 的一个纤维表示是比较平凡的. 一个获得 A 的纤维表示的更重要的途径为由 A 上一个所谓等价关系的似相等关系而得到. A 上的一个*等价关系* $\sim$ 为一个二元关系, 它是*自反的* ($x \sim x$ 对于每一个 $x \in A$ 成立),*对称的*($x \sim y \Rightarrow y \sim x$), 及*可递的*($x \sim y$ 与 $y \sim z \Rightarrow x \sim z$). A 的每一个纤维表示 $\mathfrak{F}$ 皆生成一个关系 $\sim$. 这只要规定 $x \sim y$ 当且仅当 x 与 y 属于同一个纤维, 则显然 $\sim$ 是一个等价关系. 这一节所要建立的最重要的事实为其逆.

定理 12.1 *A 上每一个等价关系 $\sim$ 为一个纤维表示的等价关系.*

证明 我们显然必须定义 $\overline{x}$ 为等价于 x 的元素 y 的集合 $\overline{x} = \{y : y \sim x\}$, 及我们的问题为证明由这一途径得到 A 的所有子集系 $\mathfrak{F}$ 为一个纤维表示.

自反, 对称与可递律变成了

$$x \in \overline{x}, \quad x \in \overline{y} \Rightarrow y \in \overline{x}, \quad \text{及} \quad x \in \overline{y}, \quad y \in \overline{z} \Rightarrow x \in \overline{z}.$$

所以由自反性推出 $\mathfrak{F}$ 包含 A. 可递性是说, 若 $y \in \overline{z}$, 则由 $x \in \overline{y} \Rightarrow x \in \overline{z}$; 即若 $y \in \overline{z}$, 则 $\overline{y} \subset \overline{z}$. 但是, 我们由对称性亦有若 $y \in \overline{z}$, 则 $z \in \overline{y}$, 所以 $\overline{z} \subset \overline{y}$. 因此 $y \in \overline{z}$ 推出 $\overline{y} = \overline{z}$. 从而若这样的两个集合 $\overline{a}$ 与 $\overline{b}$ 有一个公共元素 x, 则 $\overline{a} = \overline{x} = \overline{b}$. 换言之, 若 $\overline{a}$ 不是集合 $\overline{b}$, 则 $\overline{a}$ 与 $\overline{b}$ 互不相交, 故得一个纤维表示. □

这个论证在数学中有着基本作用是基于这样的事实, 即在许多情况下, 等价性关系是最先出现的事物, 然后用它来定义分拆与函数. 我们给出两个例子.

命 $\mathbb{Z}$ 为全体整数 (正的、负的与零). 则一个分数 'm/n' 可以被考虑为一个整数的序对 $\prec m,n \succ$, 其中 $n \neq 0$. 所以所有分数的集合为 $\mathbb{Z} \times (\mathbb{Z} - \{0\})$. 两个分数 $\prec m,n \succ$ 与 $\prec p,q \succ$ 是 "相等的" 当且仅当 $mq = np$, 及经过验证, 等号为一个等价关系. 等价类 $\prec \overline{m,n} \succ$ 为一个事物, 它被取作有理数 m/n. 因此有理数系 $\mathbb{Q}$ 是 $\mathbb{Z} \times (\mathbb{Z} - \{0\})$ 的一个分拆的纤维集合.

其次, 我们选取一个固定的整数 $p \in \mathbb{Z}$, 并且由 $mEn \Leftrightarrow p$ 整除 $m - n$, 定义了一个 $\mathbb{Z}$ 上的关系 E. 则 E 为一个等价关系, 及它的等价类集合 $\mathbb{Z}_p$ 被称为整数模 p. 易见 mEn 当且仅当 m 与 n 被 p 除之后有同样的余数, 所以在这种情况下, 存在一个易于计算的函数 f, 此处 $f(m)$ 为 m 被 p 除后的余数, 这样就定义了纤维表示. 由于可能的余数集合为 $\{0, 1, \cdots, p-1\}$, 所以 $\mathbb{Z}_p$ 包含 p 个元素.

集合 A 上的一个函数可以由下面定理经过 A 的一个纤维表示写成 "因子积".

定理 12.2 *命 g 为 A 上的一个函数, 及命 $\mathfrak{F}$ 为 A 的一个纤维表示. 则 g 在 $\mathfrak{F}$ 的每一个纤维上取常数值当且仅当 $\mathfrak{F}$ 上存在一个函数 $\overline{g}$ 使 $g = \overline{g} \circ \pi$.*

证明 若 g 在 $\mathfrak{F}$ 的每一个纤维上均为常数, 则这个惟一的值与纤维的结合定义了函数 $\overline{g}$, 显然有 $g = \overline{g} \circ \pi$. 其逆是显然的. □

第一章　向量空间

多变量函数的微积分学，是读者所熟悉的单变量微积分学与向量空间理论的结合. 这种处理的适当性直接依赖于向量空间理论真正被使用的程度. 微分方程与微分几何的理论也是类似地基于微积分与向量空间理论之融合. 这种“向量微积分”及其应用构成了本书的主要内容. 为使我们的处理完全满意，我们将在开始时，花费足够的时间来学习向量空间自身. 我们将在前两章做这件事. 本章主要讲述一般向量空间，而在下一章讲述有限维空间.

我们在本章开始时引进这一主题的基本概念 —— 向量空间，向量子空间，线性组合，与线性变量 —— 然后将这些概念与几何中的直线与平面联系起来. 其次，我们建立线性变换与笛卡儿乘积向量空间最基本的形式的性质，并简要地介绍商向量空间. 这就使我们可以在第 5 节学习第一个主题，即直和分解. 本章最后将对双线性性作一些讨论.

1.1　基本概念

向量空间与子空间. 读者可能已经接触过一些向量空间的概念. 大部分初等微积分教程都讨论过几何向量，它们用选定的原点 O 出发的“箭头”来表示. 这些向量被用平行四边形法则几何地相加：向量 $\overrightarrow{OA}$(用由 O 到 A 的箭头来表示) 与向量 $\overrightarrow{OB}$ 之和是向量 $\overrightarrow{OP}$，此处 P 是以 OA 与 OB 为边的平行四边形

中 O 的对顶点 (见图 1.1). 向量也可以与数相乘: $x(\overrightarrow{OA})$ 为向量 $\overrightarrow{OB}$, 其中 B 位于过 O 与 A 的直线上, O 至 B 的距离为 O 至 A 的距离的 $|x|$ 倍, 当 x 为正时, B 与 A 在 O 的同侧, 当 x 为负时, 它们在 O 的异侧 (图 1.2). 这两种向量运算适合某些代数法则, 我们将很快在定义中来讲述. 这些法则的几何证明一般都很简略, 其中包括了较多看似有理而不是严密的论证. 例如图 1.3 所示的几何图形就是通常关于向量加法有结合性的证明实质. 在每一种情况下, 最终的向量 $\overrightarrow{OX}$ 都可以被表示为以 OA, OB 与 OC 为边的平行多面体中以 O 为起始点的对角线. 所有几何向量的集合, 再加上这两个运算及它们满足的代数法则就构成了向量空间的一个例子. 我们将在第 2 节再来讲.

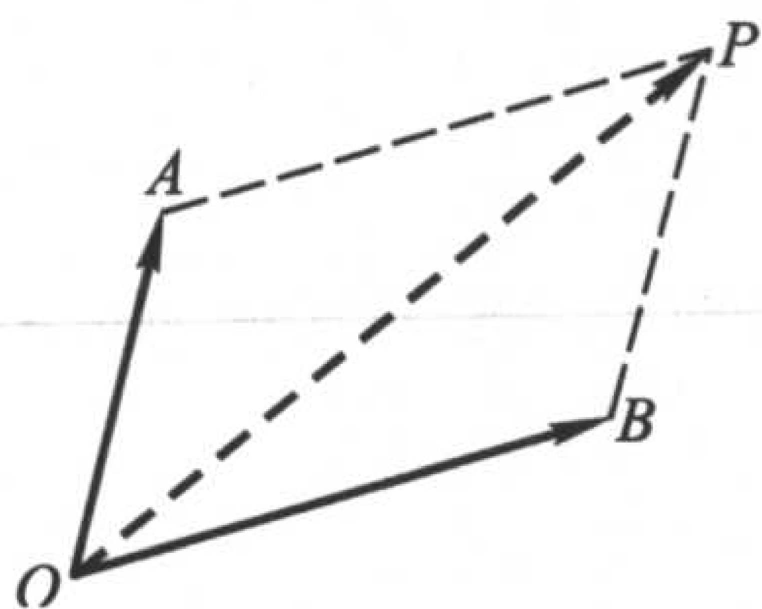


图 1.1

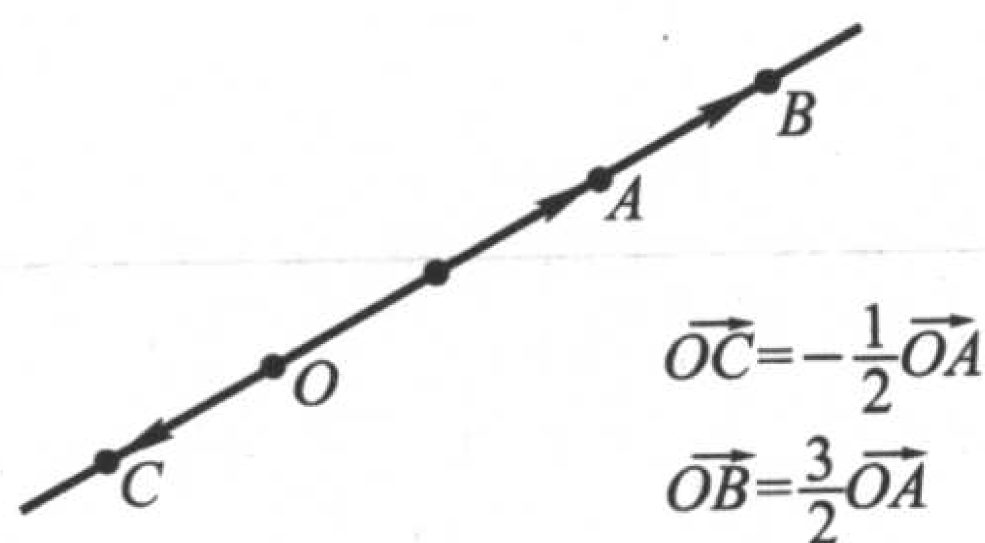
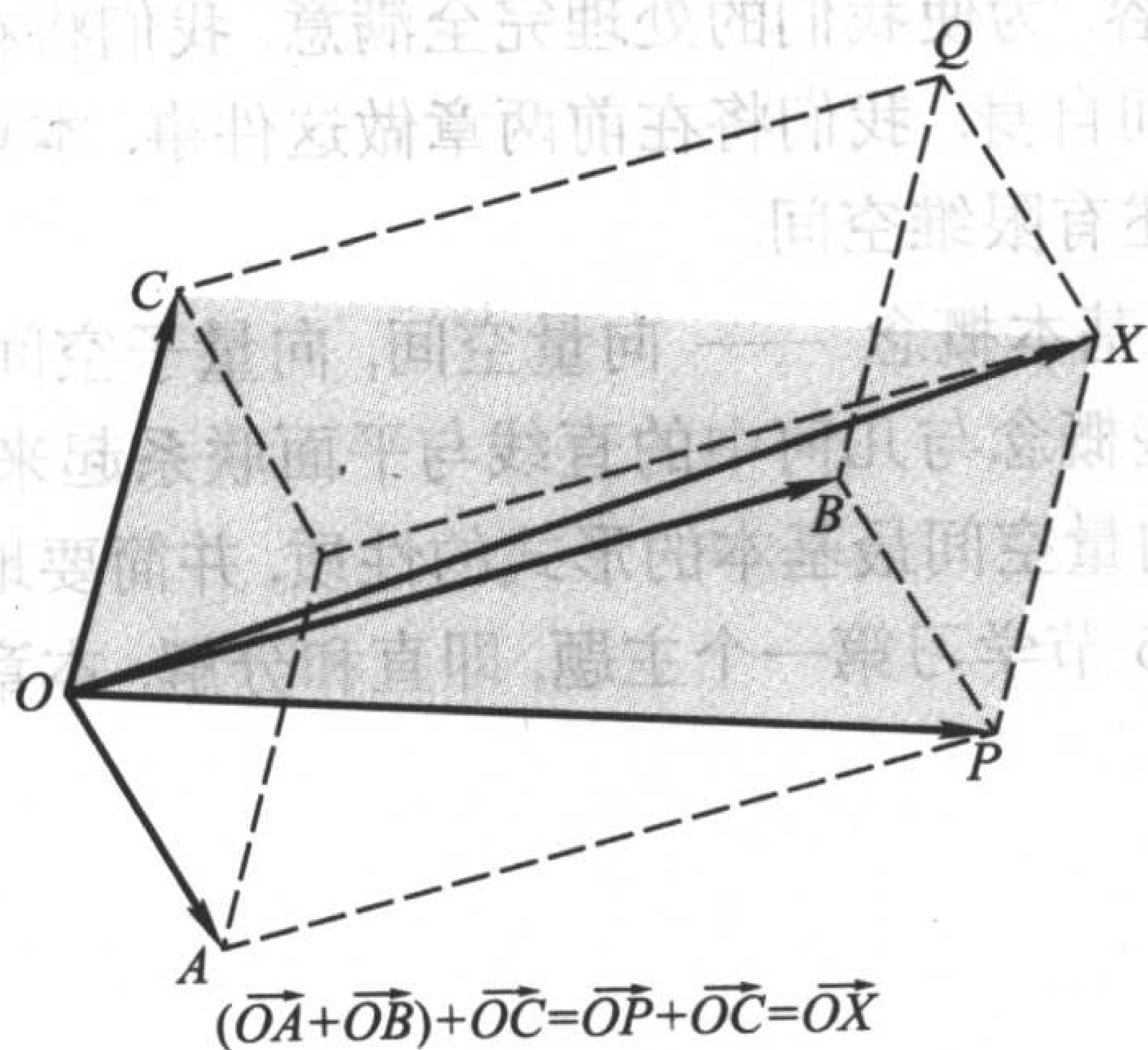


图 1.2



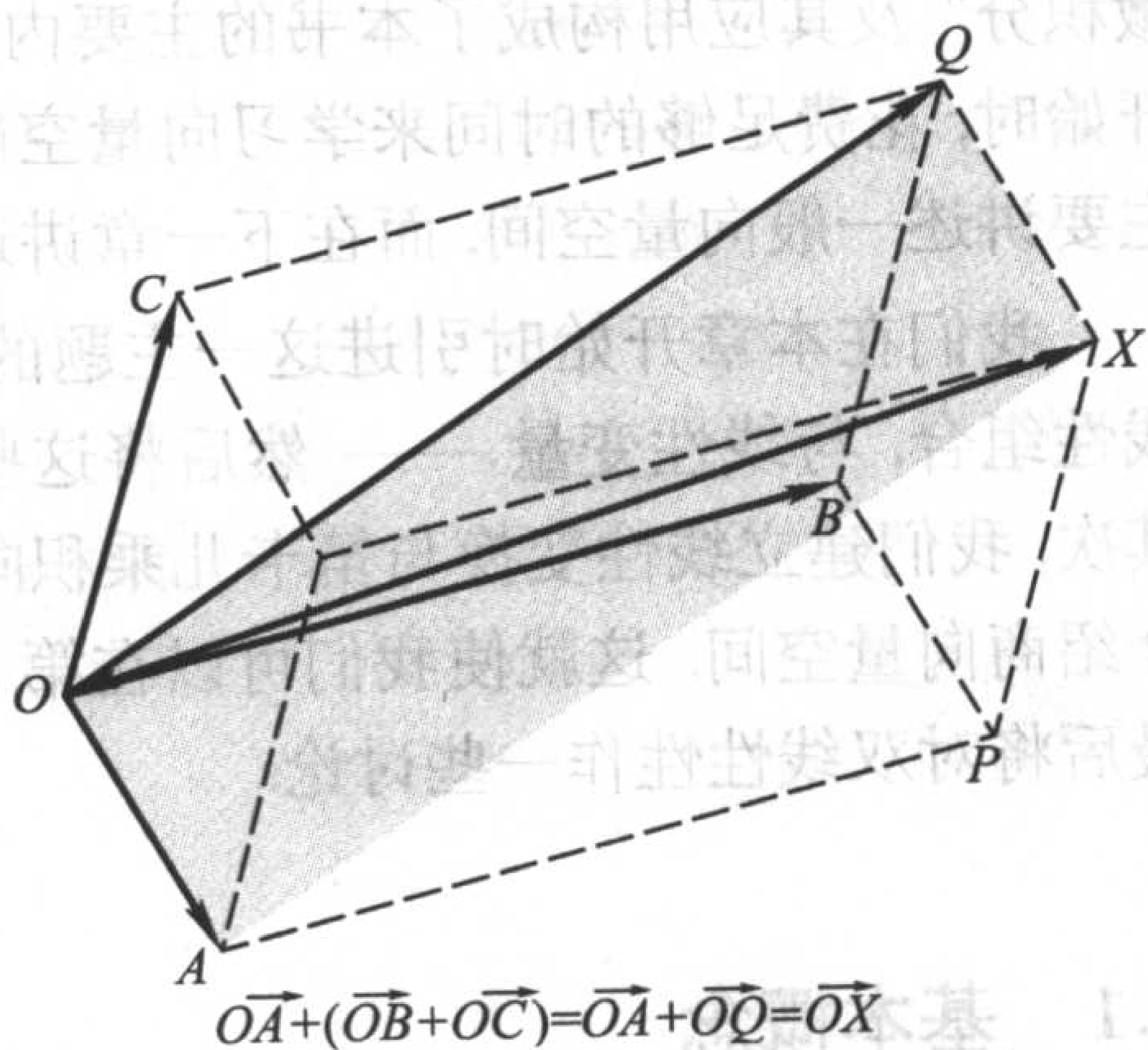


图 1.3

读者可能已经看过坐标三元组被当作向量处理. 在这个系统下, 一个三维向量是一个数的有序三元组 $\prec x_1, x_2, x_3 \succ$, 我们将它几何地设想为空间中一个点

的坐标. 加法被代数地定义为

$$\prec x_1, x_2, x_3 \succ + \prec y_1, y_2, y_3 \succ = \prec x_1 + y_1, x_2 + y_2, x_3 + y_3 \succ,$$

同样 $t \prec x_1, x_2, x_3 \succ = \prec tx_1, tx_2, tx_3 \succ$ 为数乘的定义. 由于它们已经几乎代数形式化了, 所以对于这些事物, 向量法则是很容易证明的. 因此, 所有数的有序三元组所成的集合 $\mathbb{R}^3$, 再加上这两种运算, 就是向量空间的第二个例子.

如果我们将一个有序三元组 $\prec x_1, x_2, x_3 \succ$ 设想为一个函数 $\boldsymbol{x}$, 其定义域为整数 1,2,3 构成的集合, 其中 x_i 为函数 $\boldsymbol{x}$ 在 i 之值 (见 0.8 节), 则这个向量空间暗示了一个一般的类型, 称为*函数空间*.我们将在给出定义之后, 再来讨论. 暂时, 我们仅注意三元组 $\boldsymbol{x}$ 与三元组 $\boldsymbol{y}$ 之和被定义为三元组 $\boldsymbol{z}$, 其中对于每一个 $i, z_i = x_i + y_i$.

所以, 一个向量空间为一些事物的集合, 它们之间可以彼此相加且可以乘上一个数, 并满足一些代数规则. 在本书里, 一个数常被称为一个*标量*.

定义 1.1 命 V 为一个集合. 给定一个由 $V \times V$ 至 V 的映射 $\prec \alpha, \beta \succ \mapsto \alpha + \beta$, 称为*加法*, 及一个由 $\mathbb{R} \times V$ 至 V 的映射 $\prec x, \alpha \succ \mapsto x\alpha$, 称为*标量乘积*. 如果下面诸条件满足, 则 V 称为关于这两个运算的一个*向量空间*:

A1. $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$ 对所有 $\alpha, \beta, \gamma \in V$ 成立.

A2. $\alpha + \beta = \beta + \alpha$ 对所有 α, β 成立.

A3. 存在一个元素 $0 \in V$ 使 $\alpha + 0 = \alpha$ 对所有 $\alpha \in V$ 成立.

A4. 对于任何 $\alpha \in V$, 皆存在一个 $\beta \in V$ 使 $\alpha + \beta = 0$.

S1. $(xy)\alpha = x(y\alpha)$ 对所有 $x, y \in \mathbb{R}, \alpha \in V$ 成立.

S2. $(x + y)\alpha = x\alpha + y\alpha$ 对所有 $x, y \in \mathbb{R}, \alpha \in V$ 成立.

S3. $x(\alpha + \beta) = x\alpha + x\beta$ 对所有 $x \in \mathbb{R}, \alpha, \beta \in V$ 成立.

S4. $1\alpha = \alpha$ 对所有 $\alpha \in V$ 成立.

如果在行文中, 什么样的运算是清楚的 (一般是的), 我们就简称为向量空间 V.

向量空间某些进一步的性质可以从公理中直接推出. 所以 A3 中确定的零元素是惟一的, 及对于每一个 α, A4 中的 β 也是惟一的, 记为 $-\alpha$. 还有 $0\alpha = 0, x0 = 0$, 及 $(-1)\alpha = -\alpha$. 这些基本推论都将在习题中来考虑.

向量空间的一个标准例子是一个集合 A 上所有实值函数所成的集合 $\mathbb{R}^A$, 其中两个函数的加法及一个函数与一个数的乘法都是自然运算. 这个向量空间推广了上面所说的特例 $\mathbb{R}^{\{1,2,3\}} = \mathbb{R}^3$. 记住 $\mathbb{R}^A$ 中一个函数只是某种数学事物. 我

们要说明两个这样的事物可以用一种自然的方式相加而形成第三个这样的事物,及所有这种事物所成的集合满足上述加法法则. 当然, $f+g$ 被定义为这样一个函数, 它在 a 的值为 $f(a)+g(a)$, 所以 $(f+g)(a)=f(a)+g(a)$ 对于所有 $a\in A$ 皆成立. 例如在 $\mathbb{R}^3$ 中, 我们定义 $\boldsymbol{x}+\boldsymbol{y}$ 为三元组, 对于所有 i, 它在 i 的值为 x_i+y_i. 类似地, cf 是对于所有 a, 由 $(cf)(a)=cf(a)$ 定义的函数. 由这些定义立即推出法则 A1~S4, 及关于实数系的对应代数法则. 例如方程 $(s+t)f=sf+tf$ 的意思是 $((s+t)f)(a)=(sf+tf)(a)$ 对所有 $a\in A$ 成立:

$$\begin{aligned}((s+t)f)(a)&=(s+t)(f(a))=s(f(a))+t(f(a))\\&=(sf)(a)+(tf)(a)=(sf+tf)(a).\end{aligned}$$

此处我们按次序用到了 $\mathbb{R}^A$ 中标量乘积之定义, $\mathbb{R}$ 中之分配律, $\mathbb{R}^A$ 中标量乘积之定义, 及 $\mathbb{R}^A$ 中加法的定义. 因此得 S2. 类似地, 可得其他法则.

集合 A 可以是任意东西. 若 $A=\mathbb{R}$, 则 $V=\mathbb{R}^{\mathbb{R}}$ 就是一个实变量的所有实值函数构成的向量空间. 若 $A=\mathbb{R}\times\mathbb{R}$, 则 $V=\mathbb{R}^{\mathbb{R}\times\mathbb{R}}$ 为两个实变量的所有实值函数的空间. 若 $A=\{1,2\}=\overline{2}$, 则 $V=\mathbb{R}^{\overline{2}}=\mathbb{R}^2$ 就是笛卡儿平面, 及若 $A=\{1,\cdots,n\}=\overline{n}$, 则 $V=\mathbb{R}^{\overline{n}}$ 为笛卡儿 n 维空间. 若 A 只含一个单个点, 则 $\mathbb{R}^A$ 就是 $\mathbb{R}$ 至自身的一个自然双向映射像, 当然 $\mathbb{R}$ 是关于它自身运算的一个平凡向量空间.

现在命 V 为任意向量空间, 且假定 W 是 V 的一个非空子集且在 V 的运算之下是闭的, 即若 α 与 β 在 W 中, 则 $\alpha+\beta$ 亦然. 又若 $\alpha\in W$, 则对于每一个标量 $x, x\alpha$ 亦然. 例如命 V 为闭区间 $[a,b]\subset\mathbb{R}$ 上所有实值函数所成的向量空间 $\mathbb{R}^{[a,b]}$, W 为 $[a,b]$ 上所有连续的实值函数所成的集合 $\mathfrak{C}([a,b])$. 由于当 f 与 g 为连续函数时, $f+g$ 与 cf 亦然, 所以 W 是 V 的子集, 它在 V 的运算之下是闭的. 或者命 V 为笛卡儿二维空间 $\mathbb{R}^2$, 且 W 为满足 $x_1+x_2=0$ 的序对集合 $\boldsymbol{x}=\prec x_1,x_2\succ$, 则显然在 V 的运算下, W 是闭的.

这样的一个子集 W 自身总是一个向量空间. 普遍的量化定律 A1, A2 与 S1 至 S4 在较大的集合 V 中成立, 所以它们在 W 亦成立. 由于在 W 中有某 β 及对于标量乘法, W 是闭的, 所以 $0=0\beta\in W$. 同理, 若 $\alpha\in W$, 则 $-\alpha=(-1)\alpha$ 亦然. 所以 A3 与 A4 亦成立, 从而 W 是一个向量空间. 我们证明了下面的引理.

引理 1.1 *若 W 是一个向量空间 V 的一个非空子集, 它在 V 的运算之下是闭的, 则 W 自身也是一个向量空间.*

我们称 W 为 V 的一个子空间.因此 $\mathfrak{C}([a,b])$ 是 $\mathbb{R}^{[a,b]}$ 的一个子空间, 及所有

满足 $x_1 + x_2 = 0$ 的数对 $\prec x_1, x_2 \succ$ 构成 $\mathbb{R}^2$ 的一个子空间. 子空间概念将贯穿本书之始终.

一个向量空间 $\mathbb{R}^A$ 的子空间称为一个*函数空间*. 换言之, 函数空间是一个在一个共同定义域上的实值函数集合, 其中关于加法及标量乘法是闭的.

到现在为止, 我们已经定义的应该是所谓的*实向量空间*或 $\mathbb{R}$ 上的向量空间概念. 还有类似的*复向量空间*概念, 其中标量是复数. 则定律 S1 至 S4 指的是复数相乘, 及 A 上所有复值函数所成的空间 $\mathbb{C}^A$ 就是一个标准的例子. 事实上, 如果读者已经知道一个域 F 的含义, 我们就可以给出一个单独的 F 上一个向量空间的一般定义, 而标准例子为所有由 A 至 F 的函数所成的空间 $V = F^A$. 在本书中, 如果没有特别的说明, 一个向量空间将被理解为一个实向量空间. 但是, 很多分析对复向量空间亦成立, 而且绝大部分纯代数推导对于任何标量域 F 都成立.

习题

1.1 对于几何向量, 简单描绘表示定律 S3,

$$x(\overrightarrow{OA} + \overrightarrow{OB}) = x(\overrightarrow{OA}) + x(\overrightarrow{OB})$$

的几何图形. 这里假定 $x > 1$.

1.2 试用有序三元组 $\{x_1, x_2, x_3\}$ 的显式表示, 对于 $\mathbb{R}^3$ 证明 S3.

1.3 由初等代数的简易推导即可证明 A3 中的向量 0 是惟一的. 为此假定 $0'$ 亦满足 A3. 则

$$\begin{aligned} 0' &= 0' + 0 && \text{(对 0 用 A3)} \\ &= 0 + 0' && \text{(A2)} \\ &= 0 && \text{(对 } 0' \text{用 A3).} \end{aligned}$$

试用类似的代数推导证明, 给定 α, A4 中的 β 是惟一的. 这个惟一的 β 记为 $-\alpha$.

1.4 类似地, 试证 $0\alpha = 0, x0 = 0$, 及 $(-1)\alpha = -\alpha$.

1.5 试证若 $x\alpha = 0$, 则或者 $x = 0$ 或者 $\alpha = 0$.

1.6 对于函数空间 $\mathbb{R}^A$, 试证 S1 与 S3.

1.7 给定向量空间 V 中任意向量 α, 试证所有 α 的标量积所成的集合 $\{x\alpha : x \in \mathbb{R}\}$ 是 V 的一个子空间.

1.8 给定 V 中任意两个向量 α 和 β, 试证所有形如 $x\alpha + y\beta$ 之向量所成的集合是 V 的一个子空间, 此处 x 与 y 为任意实数.

1.9 试证 $\mathbb{R}^3$ 中满足 $x_1 - x_2 + 2x_3 = 0$ 的三元组 $\boldsymbol{x}$ 的集合是一个子空间 M. 若 N 为类似的子空间 $\{\boldsymbol{x} : x_1 + x_2 + x_3 = 0\}$, 在 $M \cap N$ 中找一个非零向量 $\boldsymbol{a}$. 试证 $M \cap N$ 为 $\boldsymbol{a}$ 的所有标量积所成的集合 $\{x\boldsymbol{a} : x \in \mathbb{R}\}$.

1.10 命 A 为开区间 (0,1) 及 V 为 $\mathbb{R}^4$. 给予 (0,1) 中一个点 x, 命 V_x 为 V 中在 x 有一个导数的函数所成之集合. 试证 V_x 为 V 的子空间.

1.11 对于一个向量空间的任何子集 A 与 B, 我们由 $A + B = \{\alpha + \beta : \alpha \in A$ 与 $\beta \in B\}$ 定义集之和 $A + B$. 试证 $(A + B) + C = A + (B + C)$.

1.12 如果 $A \subset V$ 及 $X \subset \mathbb{R}$, 我们类似地定义 $XA = \{x\alpha : x \in X$ 与 $\alpha \in A\}$. 试证一个非空集合 A 为一个子空间当且仅当 $A + A = A$ 及 $\mathbb{R}A = A$.

1.13 命 V 为 $\mathbb{R}^2$, 及 M 是过原点且有斜率 k 之直线. 命 $\boldsymbol{x}$ 为 M 中任意非零向量. 试证 M 为子空间 $\mathbb{R}\boldsymbol{x} = \{t\boldsymbol{x} : t \in \mathbb{R}\}$.

1.14 试证具有同样斜率 k 的任何直线 L 皆具形式 $M + \boldsymbol{a}$, 其中 $\boldsymbol{a}$ 为某向量.

1.15 命 M 为一个向量空间 V 的一个子空间, 及 α 与 β 为 V 中任意两个向量. 给定 $A = \alpha + M$ 及 $B = \beta + M$, 试证或者 $A = B$ 或者 $A \cap B = \varnothing$. 再证明 $A + B = (\alpha + \beta) + M$.

1.16 仔细陈述并证明 "一个子空间的一个子空间是一个子空间".

1.17 试证一个向量空间的两个子空间的交本身总是一个子空间.

1.18 更一般地, 试证 V 的子空间系 $\{W_i : i \in I\}$ 的交 $W = \cap_{i \in I} W_i$ 是 V 的一个子空间.

1.19 命 V 为 $\mathbb{R}^{(0,1)}$, 及 W 为所有 V 中, 对于每一个 (0,1) 中的 x, $f'(x)$ 皆存在的函数集合. 试证 W 是子空间 V_x 所成的集合系之交, 其中 V_x 为习题 1.10 中所示的子空间.

1.20 命 V 为一个函数空间 V^A, 及对于 A 中一个点 a, 命 W_a 为满足 $f(a) = 0$ 之函数集合. 则 W_a 显然是一个子空间. 对于一个子集 $B \subset A$, 命 W_B 为满足在 B 上有 $f = 0$ 的函数 f 所成之集合. 试证 W_B 为交集 $\cap_{a \in B} W_a$.

1.21 假定 X 与 Y 为 V 的子空间, 试证若 $X + Y = V$ 及 $X \cap Y = \{0\}$, 则对于 V 中每一个向量 ζ, 皆存在一对惟一的向量 $\xi \in X$ 与 $\eta \in Y$ 满足 $\xi + \eta = \zeta$.

1.22 试证 X 与 Y 为一个向量空间 V 的子空间, 则只有当 $X \subset Y$ 或 $Y \subset X$ 时, 并 $X \cup Y$ 才可能是一个子空间.

线性组合与线性生成. 由于向量加法的交换律与结合律, 所以向量的有限集之和对于用所有可能的途径相加, 其结果均相同. 例如, 三个向量 $\alpha_a, \alpha_b, \alpha_c$ 之和可以有 12 种算法, 它们得到同样的结果:

$$(\alpha_a + \alpha_b) + \alpha_c = \alpha_c + (\alpha_a + \alpha_b) = (\alpha_c + \alpha_a) + \alpha_b = \alpha_b + (\alpha_c + \alpha_a) \quad \text{等等.}$$

因此, 如果 $I = \{a, b, c\}$ 为用到的指标集, 则记号 $\sum_{i \in I} \alpha_i$ 表示和, 虽然没有告诉我们如何相加, 但却不会混淆. 一般说来, 对于任意有限向量的有指标集合

$\{\alpha_i : i \in I\}$, 皆存在一个惟一确定的和向量 $\sum_{i\in I}\alpha_i$, 我们可以用任何次序及用任何方式组合起来相加.

指标集合 I 经常为一段整数 $\overline{n} = \{1, \cdots, n\}$. 在这种情况下, 诸向量 α_i 构成一个 n 元组 $\{\alpha_i\}_1^n$, 除了有专门的说明外, 我们将按自然次序对它们相加, 并将和记作 $\sum_{i=1}^n \alpha_i$. 注意, 它们被分组的方式仍然是任意的.

但是, 我们必须经常使用的有指标集合是无次序的. 例如, 两个变量 's' 与 't' 的次数不超过 5 的一般多项式为

$$\sum_{0\leqslant i+j\leqslant 5} c_{ij}s^i t^j,$$

单项式 $\{s^i t^j\}_{i+j\leqslant 5}$ 的有限集合并没有自然次序.

* 向量的有限集合之和与我们如何将它们相加无关之正式证明需要用归纳法. 我们给出它, 仅为有兴趣的读者阅读.

为避免看上去的烦琐, 我们由两个向量开始归纳法. 在这种情况下, 交换律 $\alpha_a + \alpha_b = \alpha_b + \alpha_a$ 给出了所有可能的和都恒等. 假定对于少于 n 个元素的指标集合, 这个结论成立. 同时考虑有 n 个元素的一个集合 $\{\alpha_i : i \in I\}$. 命 β 与 γ 为按两种方式计算出来的这些向量之和. 在 β 的计算中, 一定有最后一步加法来完成, 所以 $\beta = (\sum_{i\in J_1}\alpha_i) + (\sum_{i\in J_2}\alpha_i)$, 此处 $\{J_1, J_2\}$ 分拆 I, 我们可以记这两个部分和而不需指出它们是怎样被组成的, 这是我们的归纳法假定了用所有可能的方式相加它们都得到相同的结果.

类似地, $\gamma = (\sum_{i\in K_1}\alpha_i) + (\sum_{i\in K_2}\alpha_i)$. 现在置

$$L_{jk} = J_j \bigcap K_k \quad 与 \quad \xi_{jk} = \sum_{i\in L_{jk}} \alpha_i,$$

此处当 L_{jk} 为空集时, 我们理解为 $\xi_{jk} = 0$(见习题 1.37). 则由归纳法可知 $\sum_{i\in J_1} = \xi_{11} + \xi_{12}$, 对于其他三个和亦有类似结果. 所以

$$\beta = (\xi_{11} + \xi_{12}) + (\xi_{21} + \xi_{22}) = (\xi_{11} + \xi_{21}) + (\xi_{12} + \xi_{22}) = \gamma.$$

命题证完.*

假定 A 是向量空间 V 的一个子集合. 如果 β 是一个有限和 $\sum x_i\alpha_i$, 其中向量 α_i 均属于 A, 标量 x_i 是任意的, 则称 β 为 A 的一个线性组合.因此, 如果 A 是所有 "单项式" 的子集 $\{t^n\}_0^\infty \subset \mathbb{R}^{\mathbb{R}}$, 则一个函数 f 是 A 中函数的线性组合当且仅当 f 是一个多项式函数 $f(t) = \sum_1^n c_i t^i$. 若 A 是有限的, 则取有指标的集合

$\{\alpha_i\}$ 为整个 A 常常是很有用的, 及简单地用一个 0 系数表示任何不在和中之向量. 因此, 如果 A 是 $\mathbb{R}^{\mathbb{R}}$ 的子集合 $\{\sin t, \cos t, e^t\}$, 则我们可以将 A 当作如所示次序的一个三元组. 函数 $3\sin t - e^t = 3 \cdot \sin t + 0 \cdot \cos t + (-1)e^t$ 为三元组 A 具有系数三元组 $\prec 3, 0, -1 \succ$ 的线性组合.

现在考虑 $\mathbb{R}^3$ 中两个向量 $\prec 1,1,1 \succ$ 与 $\prec 0,1,-1 \succ$ 所有线性组合所成的集合 L. 这是所有向量 $s \prec 1,1,1 \succ + t \prec 0,1,-1 \succ = \prec s, s+t, s-t \succ$ 所成的集合, 其 s 与 t 为任意实数. 因此, $L = \{\prec s, s+t, s-t \succ : \prec s, t \succ \in \mathbb{R}^2\}$. 我们在观察后即清楚, L 在加法与标量乘法之下是闭的, 所以它是 $\mathbb{R}^3$ 的一个子空间. L 亦包含这两个给定的向量, 它们的系数对分别是 $\prec 1, 0 \succ$ 与 $\prec 0, 1 \succ$. 最后, $\mathbb{R}^3$ 的任何包含这两个给定向量的子空间 M, 亦必包含它们所有的线性组合, 从而包含 L. 即 L 是 $\mathbb{R}^3$ 中包含 $\prec 1,1,1 \succ$ 与 $\prec 0,1,-1 \succ$ 的*最小子空间*. 它被称为这两个向量的*线性生成*, 或由这两个向量*生成*的子空间. 一般言之, 我们有下面的定理.

定理 1.1 *若 A 是一个向量空间 V 的一个非空子集合, 则 A 中向量的所有线性组合所成的集合是一个子空间, 而且它是 V 中包含集合 A 的最小子空间.*

证明 首先假定 A 是有限的. 我们可以假定 A 为某种有指标的集合. 所以有某种有限指标集合使 $A = \{\alpha_i : i \in I\}$, 及 $L(A)$ 的每一元素皆有形式 $\sum_{i \in I} x_i \alpha_i$. 因此我们有

$$(\sum x_i \alpha_i) + (\sum y_i \alpha_i) = \sum (x_i + y_i) \alpha_i.$$

这是由于当重新用配对组合后, 左端就变成 $\sum_i (x_i \alpha_i + y_i \alpha_i)$, 然后由 S2 即得右端. 由 S3 及数学归纳法, 我们亦有

$$c(\sum x_i \alpha_i) = \sum (c x_i) \alpha_i.$$

所以在加法与标量乘法之下, $L(A)$ 是闭的. 从而它是一个子空集. 进而言之, $L(A)$ 包含每一个 α_i (为什么?), 所以它包含 A. 最后, 若一个子空间 W 包含 A, 则它亦包含每一个线性组合 $\sum x_i \alpha_i$, 所以它包含 $L(A)$. 因此 $L(A)$ 可以被直接地特征化为包含集合 A 的惟一确定的最小子空间.

若 A 是无限的, 则我们显然不能用单纯的有限列举办法. 但是, A 的元素的两个线性组合之和 $(\sum_1^n x_i \alpha_i) + (\sum_1^m y_j \beta_j)$ 显然是标量乘 A 的元素的一个有限和. 如果我们愿意, 我们可以将这个和改写为 $\sum_1^{n+m} x_i \alpha_i$, 此处当 $j = 1, \cdots, m$ 时, 我们置 $\beta_j = \alpha_{n+j}$ 及 $y_j = x_{n+j}$. 总之, $L(A)$ 仍然在加法与标量乘法之下是闭的, 从而它是一个子空间. □

•我们称 $L(A)$ 为 A 的线性生成. 如果 $L(A) = V$, 则我们称 A生成V; 如果 V 有一个有限生成的集合, 则称 V 是有限维的.

若 $V = \mathbb{R}^3$, 及若 δ^1, δ^2 与 δ^3 为 "轴上的单位点": $\delta^1 = \prec 1,0,0 \succ, \delta^2 = \prec 0,1,0 \succ$, 及 $\delta^3 = \prec 0,0,1 \succ$, 则 $\{\delta^i\}_1^3$ 生成 $\mathbb{R}^3$, 这是由于对于 $\mathbb{R}^3$ 中每一个点 $\boldsymbol{x}$, 我们皆有 $\boldsymbol{x} = \prec x_1, x_2, x_3 \succ = \prec x_1, 0, 0 \succ + \prec 0, x_2, 0 \succ + \prec 0, 0, x_3 \succ = x_1\delta^1 + x_2\delta^2 + x_3\delta^3 = \sum_1^3 x_i\delta^i$. 更一般地, 若 $V = \mathbb{R}^n$ 及 δ^j 为 n 元组, 其中第 j 个位置有值 1 而其他位置均取值 0, 则类似地, 我们有 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ = \sum_{i=1}^n x_i\delta^i$, 所以 $\{\delta^i\}_1^n$, 生成 $\mathbb{R}^n$. 一般说来, 一个无限集合 A 上的一个函数空间不是有限维的. 例如, $\mathfrak{C}([a,b])$ 没有有限的生成集合. 这个事实是对的, 但并不显然.

习题

1.23 给定 $\alpha = \prec 1,1,1 \succ, \beta = \prec 0,1,-1 \succ, \gamma = \prec 2,0,1 \succ$, 试计算线性组合 $\alpha + \beta + \gamma, 3\alpha - 2\beta + \gamma, x\alpha + y\beta + z\gamma$, 并寻找 x, y 与 z 使 $x\alpha + y\beta + z\gamma = \prec 0,0,1 \succ = \delta^3$. 对于 δ^1 与 δ^2, 试做同样的事.

1.24 给定 $\alpha = \prec 1,1,1 \succ, \beta = \prec 0,1,-1 \succ, \gamma = \prec 1,0,2 \succ$, 试证 α, β, γ 中, 每一个都是其余两个的线性组合. 并证明不可能找到满足 $x\alpha + y\beta + z\gamma = \delta^1$ 的系数 x, y 与 z.

1.25 (a) 试求集合 $A = \prec t, t^2-1, t^2+1 \succ$ 与系数三元组 $\prec 2,-1,1 \succ$ 的线性组合, 并对 $\prec 0,1,1 \succ$ 做同样的事.

(b) 试求系数三元组使与三元组 A 的线性组合为 $(t+1)^2$, 并对 1 做同样的事.

(c) 试证任何次数 $\leqslant 2$ 的多项式事实上是 A 的一个线性组合.

1.26 试找出 $\{e^t, e^{-t}\} \subset \mathbb{R}^{\mathbb{R}}$ 满足 $f(0) = 1$ 及 $f'(0) = 2$ 的线性组合 f.

1.27 试找出 $\sin x, \cos x$ 与 e^x 的一个线性组合 f, 它满足 $f(0) = 0, f'(0) = 1$ 及 $f''(0) = 1$.

1.28 假定 $a\sin x + b\cos x + ce^x$ 是零函数. 试证 $a = b = c = 0$.

1.29 试证 $\prec 1,1 \succ$ 与 $\prec 1,2 \succ$ 生成 $\mathbb{R}^2$.

1.30 试证子空间 $M = \{\boldsymbol{x} : x_1 + x_2 = 0\} \subset \mathbb{R}^2$ 可以被一个向量生成.

1.31 命 M 为 $\mathbb{R}^3$ 的子空间 $\{\boldsymbol{x} : x_1 - x_2 + 2x_3 = 0\}$. 试求 M 中两个向量 $\boldsymbol{a}$ 与 $\boldsymbol{b}$, 其中任何一个都不是另一个标量倍数. 试证 M 是 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 的线性生成.

1.32 试求 $\mathbb{R}^3$ 中 $\prec 1,1,1 \succ$ 与 $\prec 0,1,-1 \succ$ 的线性生成与坐标子空间 $x_2 = 0$ 的交集, 并将这个交集表示为一个线性生成.

1.33 将坐标空间换成

$$M = \{\boldsymbol{x} : x_1 + x_2 = 0\}$$

再做一下上面的习题.

1.34 由定理 1.1 可知一个向量空间 V 的任意子集合 A 的线性生成 $L(A)$ 有下面两个性质:
(i) $L(A)$ 是 V 的包含 A 的一个子空集;
(ii) 若 M 为任何包括 A 的子空间, 则 $L(A) \subset M$.

试仅使用 (i) 与 (ii) 证明
(a) $A \subset B \Rightarrow L(A) \subset L(B)$;
(b) (b) $L(L(A)) = L(A)$.

1.35 试证

(a) 若 M 与 N 为 V 的子空间, 则 $M + N$ 亦然;

(b) 对于任意子集合 $A, B \subset V$ 有 $L(A \cup B) = L(A) + L(B)$.

1.36 回忆 (习题 1.18) 任何子空间系的交集仍为一个子空间. 试证一个向量空间 V 的一个子集合 A 的线性生成 $L(A)$ 是 V 的所有包含有 A 的子空间之交集. 这个另外的特征化有时被取作线性生成的定义.

1.37 按惯例, 向量的空集元素之和被取作零向量. 如果定理 1.1 是严格正确的, 则这是必要的. 为什么? 对于上面一个问题呢?

线性变换 一般的函数空间 $\mathbb{R}^A$ 与 $\mathbb{R}^{[a,b]}$ 的子空间 $\mathfrak{C}([a,b])$ 都有这样的性质 : 除了在向量运算之下为闭的之外, 它们在两个函数乘积运算之下亦是闭的. 即两个函数的逐点乘积仍是一个函数 $[(fg)(a) = f(a)g(a)]$, 及两个连续函数之积仍为连续的. 关于加, 乘与标量积这三种运算, $\mathbb{R}^A$ 与 $\mathfrak{C}([a,b])$ 都是代数的例子. 如果读者注意到这个附加的运算, 他可能感到奇怪, 至少在函数空间的讨论中, 为什么我们讨厌向量空间的概念. 为什么不学习所有这三个运算? 回答是向量运算恰好是这种运算, 它被函数集合的许多最重要的映射所 "保持". 例如, 由 $T(f) = \int_a^b f(t)dt$ 所定义的 $T : \mathfrak{C}([a,b]) \to \mathbb{R}$. 则积分学的定律说 $T(f+g) = T(f) + T(g)$ 及 $T(cf) = cT(f)$. 因此 T "保持" 了向量运算. 由于 "加法接着 T" 等于 "T 接着加法", 所以我们也可以说, T 与向量运算是 "交换的". 但是 T 不保持乘法 : 一般说来, $T(fg) = T(f) \cdot T(g)$ 不成立.

另一个例子是由 $y_1 = 2x_1 - x_2 + x_3, y_2 = x_1 + 3x_2 - 5x_3$ 定义的由 $\mathbb{R}^3$ 至 $\mathbb{R}^2$ 的映射 $T : \boldsymbol{x} \mapsto \boldsymbol{y}$. 我们仍然可以验证 $T(\boldsymbol{x}+\boldsymbol{y}) = T(\boldsymbol{x}) + T(\boldsymbol{y})$ 与 $T(c\boldsymbol{x}) = cT(\boldsymbol{x})$. 线性方程组可解性理论本质上就是这样映射 T 的理论; 所以我们有了另一类重要映射, 它保持向量运算 (但不保持乘积运算).

这些注记建议我们学习一些向量空间, 从而我们可以知道保持向量运算的映射. 这种映射被称为线性变换.

定义 1.2 若 V 与 W 为向量空间, 则若一个映射 $T:V \to W$ 满足: $T(\alpha+\beta) = T(\alpha)+T(\beta)$, 对所有 α, $\beta \in V$, 及 $T(x\alpha) = xT(\alpha)$ 对所有 $\alpha \in V, x \in \mathbb{R}$ 成立, 就称 T 为一个线性变换或一个线性映射.

T 满足的这两个条件可以被联合成一个单个方程

$$T(x\alpha + y\beta) = xT(\alpha) + yT(\beta), \text{ 对所有}\alpha, \beta \in V\text{及所有}x, y \in \mathbb{R}\text{成立}.$$

进而言之, 这个方程还可以由归纳法推广至任意有限和的情况, 所以若 T 是线性的, 则

$$T\left(\sum_{i\in I} x_i\alpha_i\right) = \sum_{i\in I} x_i T(\alpha_i)$$

对于所有线性组合 $\sum x_i\alpha_i$ 成立. 例如 $\int_a^b(\sum_1^n c_i f_i) = \sum_1^n c_i \int_a^b f_i$.

习题

1.38 试证由 $\mathbb{R}$ 至 $\mathbb{R}$ 最一般的线性变换为乘以一个常数.

1.39 对于 V 中一个固定的 α, 由 $\mathbb{R}$ 至 V 的映射 $x \mapsto x\alpha$ 是线性的. 为什么?

1.40 当 x 固定时, 为什么这对于 $\alpha \mapsto x\alpha$ 成立?

1.41 试证每一个由 $\mathbb{R}$ 至 V 的线性变换皆取形式 $x \mapsto x\alpha$, 其中 α 为 V 中一个固定向量.

1.42 试证由 $\mathbb{R}^2$ 至 V 的每一个线性变换皆有形式 $\prec x_1, x_2 \succ \mapsto x_1\alpha_1 + x_2\alpha_2$, 其中 α_1 与 α_2 为 V 中一对固定的向量. 这个映射的值是什么?

1.43 试证由 $\mathfrak{C}([a,b])$ 至 $\mathbb{R}$ 的映射 $f \mapsto \int_a^b f(t)$ 不保持乘积.

1.44 命 g 为 $\mathbb{R}^A$ 中任意固定函数, 试证由 $T(f) = gf$ 定义的映射 $T: \mathbb{R}^A \to \mathbb{R}^A$ 是线性的.

1.45 命 φ 为由一个集合 A 至一个集合 B 的任意映射. 试证由 φ 的合成是一个由 $\mathbb{R}^B$ 至 $\mathbb{R}^A$ 的线性映射, 即证由 $T(f) = f \circ \varphi$ 定义的映射 $T: \mathbb{R}^B \to \mathbb{R}^A$ 是线性的.

为了获得一些例子, 我们将找出所有以 $\mathbb{R}^n$ 为定义空间的线性变换. 我们可以考虑一个这种变换作为很好的开端. 假定我们在 $\mathbb{R}$ 上所有实值函数空间 $\mathbb{R}^{\mathbb{R}}$ 中选取一个固定的函数三元组 $\{f_i\}_1^3$, 其中 $f_1(t) = \sin t, f_2(t) = \cos t$, 及 $f_3(t) = e^t = \exp(t)$. 则对于 $\mathbb{R}^3$ 中每一个三元数组 $\boldsymbol{x} = \{x_i\}_1^3$, 我们皆有一个以 $\{x_i\}$ 为系数的线性组合 $\sum_{i=1}^3 x_i f_i(t) = x_1\sin t + x_2\cos t + x_3 e^t$. 不同的系数三元组给出不同的函数, 从而映射 $\boldsymbol{x} \mapsto \sum_{i=1}^3 x_i f_i = x_1\sin + x_2\cos + x_3\exp$ 是一个由 $\mathbb{R}^3$ 至 $\mathbb{R}^{\mathbb{R}}$ 的映射. 它显然是线性的. 如果我们称这个映射为 T, 则我们可以由

T 恢复决定的函数三元组为 $\mathbb{R}^3$ 中 "单位点"δ^i 的像; $T(\delta^j)=\sum \delta_i^j f_i = f_j$, 所以 $T(\delta^1)=\sin$, $T(\delta^2)=\cos$, 及 $T(\delta^3)=\exp$. 我们来检验由 $\mathbb{R}^3$ 至 $\mathbb{R}^{\mathbb{R}}$ 的每一个线性映射都具有这种形式.

在下面的定理中, $\{\delta^i\}_1^n$ 为以前定义过的 $\mathbb{R}^n$ 的生成集合, 所以对于 $\mathbb{R}^n$ 中每一个 n 元组 $\boldsymbol{x}=\prec x_1,\cdots,x_n\succ$ 皆有 $\boldsymbol{x}=\sum_1^n x_i\delta^i$.

定理 1.2 若 $\{\beta_j\}_1^n$ 为一个向量空间 W 中任意固定的向量 n 元组, 则 "线性组合映射" $\boldsymbol{x}\mapsto\sum_1^n x_i\beta_i$ 为一个由 $\mathbb{R}^n$ 至 W 的线性变换 T, 及对于 $j=1,\cdots,n$, 有 $T(\delta^i)=\beta_j$. 反之, 若 T 为任意由 $\mathbb{R}^n$ 至 W 的线性映射, 及若当 $j=1,\cdots,n$ 时, 我们置 $\beta_j=T(\delta^j)$, 则 T 为线性组合映射 $\boldsymbol{x}\mapsto\sum_1^n x_i\beta_i$.

证明 用定理 1.1 中证明 $L(A)$ 为一个子空间完全同样的论证方法, 可以导出线性组合映射 T 的线性性. 因此

$$\begin{aligned}T(\boldsymbol{x}+\boldsymbol{y})&=\sum_1^n(x_i+y_i)\beta_i=\sum_1^n(x_i\beta_i+y_i\beta_i)\\&=\sum_1^n x_i\beta_i+\sum_1^n y_i\beta_i=T(\boldsymbol{x})+T(\boldsymbol{y}),\end{aligned}$$

与

$$T(s\boldsymbol{x})=\sum_1^n(sx_i)\beta_i=\sum_1^n s(x_i\beta_i)=s\sum_1^n x_i\beta_i=sT(\boldsymbol{x}).$$

由于 $\delta_j^j=1$ 及 $\delta_i^j=0(i\neq j)$, 所以 $T(\delta^j)=\sum_{i=1}^n\delta_i^j\beta_i=\beta_j$.

反之, 若 $T:\mathbb{R}^n\to W$ 是线性的, 及若对于所有 j, 我们置 $\beta_j=T(\delta^j)$, 则对于 $\mathbb{R}^n$ 中任意 $\boldsymbol{x}=\prec x_1,\cdots,x_n\succ$, 我们有 $T(\boldsymbol{x})=T(\sum_1^n x_i\delta^i)=\sum_1^n x_iT(\delta^i)=\sum_1^n x_i\beta_i$. 因此 T 为映射 $\boldsymbol{x}\mapsto\sum_1^n x_i\beta_i$. □

这是一个看似简单但非常重要的定理, 读者应该牢记于心. 为此, 我们将发明某些前三章一直都要用的术语. 若 $\boldsymbol{\alpha}=\{\alpha_1,\cdots,\alpha_n\}$ 为一个向量空间 W 中的一个向量 n 元组, 命 $L_{\boldsymbol{\alpha}}$ 为由 $\mathbb{R}^n$ 至 W 的对应线性组合映射 $\boldsymbol{x}\mapsto\sum_1^n x_i\alpha_i$. 注意 n 元组 $\boldsymbol{\alpha}$ 自身是 W^n 的一个元素. 若 T 为任意由 $\mathbb{R}^n$ 至 W 的线性映射, 我们将称 n 元组 $\{T(\delta^i)\}_1^n$ 为 T 的骨架.用这些术语, 定理可以被重述如下:

定理 1.2′ 对于 W^n 中每一个 n 元组 $\boldsymbol{\alpha}$, 映射 $L_{\boldsymbol{\alpha}}:\mathbb{R}^n\to W$ 是线性的, 其骨架为 $\boldsymbol{\alpha}$. 反之, 若 T 是任意由 $\mathbb{R}^n$ 至 W 的线性映射, 则 $T=L_{\boldsymbol{\beta}}$, 此处 $\boldsymbol{\beta}$ 是 T 的骨架.

定理 1.2″ 映射 $\alpha \mapsto L_\alpha$ 是 W^n 至所有 $\mathbb{R}^n$ 至 W 的线性映射 T 的集合的一个双射映射, 及 $T \mapsto$ 骨架 (T) 是它的逆映射.

由一个向量空间 V 至标量场 $\mathbb{R}$ 的线性变换称为 V 上的一个线性泛函.因此 $f \mapsto \int_a^b f(t)dt$ 是 $V = \mathfrak{C}([a,b])$ 上的一个线性泛函. 上面定理对于一个线性泛函 F 来说特别简单: 这是由于 $W = \mathbb{R}$, 所以在 F 的骨架中, 每一个向量只是一个数 b_i, 从而骨架 $\{b_i\}_1^n$, 为 $\mathbb{R}^n$ 中一个元素. 在这种情况下, 我们将记 $F(\boldsymbol{x}) = \sum_1^n b_i x_i$, 即将数值系数 '$b_i$' 置于变量 '$x_i$' 之前. 因此, $F(\boldsymbol{x}) = 3x_1 - x_2 + 4x_3$ 为 $\mathbb{R}^3$ 上具有骨架 $\prec 3, -1, 4 \succ$ 的线性泛函. $\mathbb{R}^n$ 上所有线性泛函的集合与 $\mathbb{R}^n$ 本身有一个自然的一一对应. 我们由对于所有 i, $b_i = F(\delta^i)$, 从 F 得到 $\boldsymbol{b}$; 由对于所有 $\mathbb{R}^n$ 中的 $\boldsymbol{x}, F(\boldsymbol{x}) = \sum b_i x_i$, 从 $\boldsymbol{b}$ 得到 F.

其次, 我们考虑 T 的值域空间为笛卡儿空间 $\mathbb{R}^m$ 之情况, 为了在我们心中清楚这两个空间, 我们暂时取定义域为 $\mathbb{R}^3$. 现在在 T 的骨架中每一个向量 $\beta_i = T(\delta^i)$ 为一个数的 m 元组. 我们如果将这个 m 元组画成一个数的列, 则三个 m 元组 β_i 就可以被画成一个数的矩形阵, 它包含三个列, 其中每一列含有 m 个数. 命 t_{ij} 表示第 j 列中的第 i 个数, 则数的二重有指标集合 $\{t_{ij}\}$ 就被称为变换 T 的矩阵. 因为画出来的矩形阵有 m 行及 3 列, 所以我们称它为一个 m 及 3 (或 $m \times 3$) 矩阵. 由于矩阵的列构成 T 的骨架, 所以矩阵惟一地决定了 T. 恒等式 $T(\boldsymbol{x}) = \sum_1^3 x_j T(\delta^j) = \sum_1^3 x_j \beta_j$ 可使 m 元组 $T(\boldsymbol{x})$ 由 $\boldsymbol{x}$ 及矩阵 $\{t_{ij}\}$ 被显式地计算出来. 列 m 元组 β_j 与标量 x_j 相乘, 然后将三列在第 i 行处相加如下:

$$\underset{\boldsymbol{y}}{\begin{bmatrix} \vdots \\ y_i \\ \vdots \end{bmatrix}} = x_1 \underset{\beta_1}{\begin{bmatrix} \vdots \\ t_{i1} \\ \vdots \end{bmatrix}} + x_2 \underset{\beta_2}{\begin{bmatrix} \vdots \\ t_{i2} \\ \vdots \end{bmatrix}} + x_3 \underset{\beta_3}{\begin{bmatrix} \vdots \\ t_{i3} \\ \vdots \end{bmatrix}}.$$

由于 t_{ij} 为 m 元组 β_j 的第 i 个数, 所以 m 元组 $\sum_{j=1}^3 x_j \beta_j$ 的第 i 个数是 $\sum_{j=1}^3 x_j t_{ij}$, 即若我们命 $\boldsymbol{y}$ 表示 m 元组 $T(\boldsymbol{x})$, 则

$$y_i = \sum_{j=1}^3 t_{ij} x_j, \quad i = 1, \cdots, m,$$

及这 m 个标量方程的集合等价一个向量方程 $\boldsymbol{y} = T(\boldsymbol{x})$.

在上面的讨论中, 我们可以将 3 换成 n, 除了图形外, 其他均不改变, 所以我们得到特殊化的定理 1.2 如下.

定理 1.3 每一个由 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射 T 决定了 $m \times n$ 矩阵 $\boldsymbol{t} = \{t_{ij}\}$，它以 T 的骨架为列. 方程 $\boldsymbol{y} = T(\boldsymbol{x})$ 的线性组合形式表达式等价于 m 个标量方程

$$y_i = \sum_{j=1}^{n} t_{ij} x_j, \quad i = 1, \cdots, m.$$

反之, 每一个 $m \times n$ 矩阵 $\boldsymbol{t}$ 决定线性组合映射, 它以 $\boldsymbol{t}$ 的列为其骨架. 所以映射 $\boldsymbol{t} \mapsto T$ 是由所有 $m \times n$ 矩阵所成之集合与所有由 $\mathbb{R}^n$ 至 $\mathbb{R}^m$ 的线性映射所成的集合间的一个双射映射.

$\mathbb{R}^n$ 上的一个线性泛函 F 是一个由 $\mathbb{R}^n$ 至 $\mathbb{R}^1$ 的线性映射, 所以它必须被一个 $1 \times n$ 矩阵来表示, 即 $\mathbb{R}^n$ 中的 n 元组 $\boldsymbol{b}$, 它是 F 的骨架, 被看成一个 1 行与 n 列的矩阵.

作为线性映射最后一个例子, 我们考察定义在任何函数空间上的特殊线性泛函的一个重要类, 即所谓坐标泛函.若 $V = \mathbb{R}^I$ 及 $i \in I$, 则第 i 个坐标泛函 π_i 简单地为在 i 的赋值, 所以 $\pi_i(f) = f(i)$. 这些泛函显然是线性的. 事实上, 函数上的向量运算是使它们是线性的而被定义的; 由于 $sf + tg$ 被定义为那个函数, 对于所有 i, 它在 i 之值为 $sf(i) + tg(i)$, 所以我们注意到由定义, $sf + tg$ 就是那个使 $\pi_i(sf + tg) = s\pi_i(f) + t\pi_i(g)$ 对所有 i 成立的函数!

若 V 为 $\mathbb{R}^n$, 则 π_j 为映射 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ \mapsto x_j$. 则由定理可知在这种情况下, π_j 必须具有形式 $\pi_j(\boldsymbol{x}) = \sum_1^n b_i x_i$, 其中 $\boldsymbol{b}$ 为某 n 元组. $\boldsymbol{b}$ 是什么?

线性性之性质的一般形式, $T(\sum x_i \alpha_i) = \sum x_i T(\alpha_i)$. 表明 T 与 T^{-1} 均将子空间映射为子空间.

定理 1.4 若 $T: V \to W$ 是线性的, 则任何子集 $A \subset V$ 的线性生成之 T 像是 A 的 T 像的线性生成: $T[L(A)] = L[T(A)]$. 特别, 若 A 是一个子空间, 则 $T[A]$ 亦然, 进而言之, 若 Y 是 W 的一个子空间, 则 $T^{-1}[Y]$ 为 V 的一个子空间.

证明 按照公式 $T(\sum x_i \alpha_i) = \sum x_i T(\alpha_i), W$ 中的一个向量为 A 上一个线性组合的 T 像当且仅当它是 $T[A]$ 上的一个线性组合. 即 $T[L(A)] = L[T(A)]$. 如果 A 是一个子空间, 则 $A = L(A)$ 及 $T[A] = L(T[A])$ 为 W 的一个子空间. 最后, 若 Y 是 W 的一个子空间及 $\{\alpha_i\} \subset T^{-1}[Y]$, 则 $T(\sum x_i \alpha_i) = \sum x_i T(\alpha_i) \in L(Y) = Y$. 因此 $\sum x_i \alpha_i \in T^{-1}[Y]$ 及 $T^{-1}[Y]$ 是它自身的线性生成. □

子空间 $T^{-1}(0) = \{\alpha \in V : T(\alpha) = 0\}$ 称为 T 的零空间,或核,记为 $N(T)$ 或

$\mathfrak{N}(T)$. T 的值域为 W 的子空间 $T[V]$. 我们将它记为 $R(T)$ 或 $\mathfrak{R}(T)$.

引理 1.2 一个线性映射 T 是单射映射当且仅当它的零空间为 $\{0\}$.

证明 若 T 是单射映射及 $\alpha \neq 0$, 则 $T(\alpha) \neq T(0) = 0$, 从而零空间只包含 0. 另一方面, 若 $N(T) = \{0\}$. 则当 $\alpha \neq \beta$ 时, 我们有 $\alpha - \beta \neq 0, T(\alpha) - T(\beta) = T(\alpha - \beta) \neq 0$ 及 $T(\alpha) \neq T(\beta)$. 这表示 T 是单射映射. □

若一个线性映射 $T: V \to W$ 是双射的, 则称为*同构*.两个向量空间 V 与 W 是同构的当且仅当它们之间存在一个同构.

例如, 映射 $\prec c_1, \cdots, c_n \succ \to \sum_0^{n-1} c_{i+1} x^i$ 是一个 $\mathbb{R}^n$ 与所有次数小于 n 的多项式向量空间的同构.

同构空间 "具有同样的形式", 作为抽象的向量空间它们是恒等的, 即我们不能仅仅由它们具有或不具有向量性质来区别它们.

当一个线性变换是以 V 到其自身时, 一些特别的事情就可能发生. 一个可能性为 T 可能将一个向量 α 本质上映射到自身, 即对于 $\mathbb{R}$ 中某 x 使 $T(\alpha) = x\alpha$. 在这种情况下, α 称为一个*特征向量*(真向量、本征向量) 及 x 为它对应的*特征值*.

习题

1.46 在习题 1.45 的情况下, 试由证明以下事实来证明若 φ 为一个双射映射, 则 T 是一个同构:

(a) φ 为单射映射 $\Rightarrow T$ 为满射映射,

(b) φ 为满射映射 $\Rightarrow T$ 为单射映射.

1.47 试求出 $\mathbb{R}^2$ 上的线性泛函 l 满足 $l(\prec 1,1 \succ) = 0$ 及 $l(\prec 1,2 \succ) = 1$, 即在 $\mathbb{R}^2$ 中找出 $\boldsymbol{b} = \prec b_1, b_2 \succ$ 使 l 为线性组合映射

$$\boldsymbol{x} \to b_1 x_1 + b_2 x_2.$$

1.48 关于 $l(\prec 2,1 \succ) = -3$ 与 $l(\prec 1,2 \succ) = 4$ 做同样的事.

1.49 试求满足 $T(\prec 1,1 \succ) = t^2$ 与 $T(\prec 1,2 \succ) = t^3$ 的线性变换 $T: \mathbb{R}^2 \to \mathbb{R}^{\mathbb{R}}$, 即找出函数 $f_1(t)$ 与 $f_2(t)$ 使 T 为线性组合映射 $\boldsymbol{x} \to x_1 f_1 + x_2 f_2$.

1.50 命 T 为由 $\mathbb{R}^2$ 至 $\mathbb{R}^3$ 满足 $T(\delta^1) = \prec 2, -1, 1 \succ$, $T(\delta^2) = \prec 1, 0, 3 \succ$ 的线性映射. 试用标准的矩阵形式写出 T 的矩阵. 决定 δ^1 是否在 T 的值域之中.

1.51 命 T 为由 $\mathbb{R}^3$ 至 $\mathbb{R}^3$ 的线性映射, 其矩阵为

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 0 & -1 \\ 3 & -1 & 1 \end{bmatrix}$$

当 $\boldsymbol{x} = \prec 1,1,0 \succ$ 时, 试求 $T(\boldsymbol{x})$; 对于 $\boldsymbol{x} = \prec 3,-2,1 \succ$ 时, 试做同样的事.

1.52 命 M 为 $\prec 1,-1,0 \succ$ 与 $\prec 0,1,1 \succ$ 的线性生成. 试用寻找两个生成向量的办法找出子空间 $T[M]$, 其中 T 由上一习题所示.

1.53 命 T 为由 $\mathbb{R}^2$ 到自身的映射 $\prec x,y \succ \to \prec x+2y, y \succ$. 试证 T 是一个线性组合映射, 并用标准形式写出它的矩阵.

1.54 对于由 $\mathbb{R}^3$ 至自身的 $T: \prec x,y,z \succ \to \prec x-z, x+z, y \succ$ 做同样的事.

1.55 试找出由 $\mathbb{R}^3$ 至自身的线性变换 T, 它的值域空间为 $\prec 1,-1,0 \succ$ 与 $\prec -1,0,2 \succ$ 的生成.

1.56 试在 $\mathbb{R}^4$ 上找出两个线性泛函, 它们的零空间之交为 $\prec 1,1,1,1 \succ$ 与 $\prec 1,0,-1,0 \succ$ 的线性生成. 现在你已经掌握了一个线性变换,其零空间为上面的生成, 它是什么?

1.57 命 $V = \mathfrak{C}([a,b])$ 为 $[a,b]$ 上连续实值函数所成之空间, 我们亦将它记为 $\mathfrak{C}^0([a,b])$, 又命 $W = \mathfrak{C}^1([a,b])$ 为其中有连续一次导数的函数集合. 命 $D: W \to V$ 为微分 $(Df = f')$, 并由 $T(f) = F$ 定义 V 上的 T, 此处 $F(x) = \int_a^x f(x)dt$. 试由微积分的适当定理证明 D 与 T 是线性的, T 映射至 W, 及 D 为 T 的一个左逆元 ($D \circ T$ 为 V 上的恒等映射.)

1.58 在上面习题中, 将 T 的值域与 D 的零空间等同起来. 我们知 D 是满射映射及 T 是单射映射. 为什么?

1.59 命 V 为 $\sin x$ 与 $\cos x$ 的线性生成. 则微分运算 D 是由 V 至 V 的一个线性变换. 试证 D 为 V 至 V 的一个同构, 并证明在 V 上有 $D^2 = -I$.

1.60 (a) 正如读者可能猜想的, $\mathfrak{C}^3(\mathbb{R})$ 为 $\mathbb{R}$ 上有包含 3 次在内的导数之实值函数所成之集合. 试证 $f \to f'''$ 为一个由 $\mathfrak{C}^3(\mathbb{R})$ 至 $\mathfrak{C}(\mathbb{R})$ 的满射线性映射 T.

(b) 对于 $\mathbb{R}$ 中任意固定的 a, 试证 $f \to \prec f(a), f'(a), f''(a) \succ$ 为一个由零空间 $N(T)$ 至 $\mathbb{R}^3$ 的同构 [提示: 应用带余项的泰勒公式].

1.61 一个矩阵方程 $y_i = \sum_j t_{ij}x_j, i = 1, \cdots, m$ 的积分类似为方程

$$g(s) = \int_0^1 K(s,t)f(t)dt, \quad s \in [0,1].$$

假定 $K(s,t)$ 在方形 $[0,1]\times[0,1]$ 上定义, 而且对于每一个 s, 它是 t 的连续函数, 试验证 $f \to g$ 是由 $\mathfrak{C}([0,1])$ 到 $\mathbb{R}^{[0,1]}$ 的线性映射.

1.62 对于一个有限集合 $A=\{\alpha_i\}$, 定理 1.1 是定理 1.4 的一个推论. 为什么?

1.63 试证一个同构之逆是线性的 (从而亦是一个同构).

1.64 若 T 的矩阵为

$$\begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix},$$

试求 $T:\mathbb{R}^2\to\mathbb{R}^2$ 的特征向量与特征值. 由于每一个特征向量 α 的标量倍数 $x\alpha$ 显然仍为一个特征向量, 所以在每一个 "特征方向" 上寻求一个向量即可. 这是初等代数中的一个问题.

1.65 试寻求变换 T 的特征向量与特征值, 其中 T 的矩阵为

$$\begin{bmatrix} 1 & 1 \\ -2 & 0 \end{bmatrix},\quad \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix},\quad \begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix},\quad \begin{bmatrix} 2 & 1 \\ -4 & -2 \end{bmatrix}$$

1.66 上面两个习题中的 5 个变换, 按照它们具有的不同特征方向之个数展示出 4 种不同之类型. 可能性是什么?

1.67 命 V 为次数 $\leqslant 3$ 的多项式所成之向量空间, 及由 $f\to tf'(t)$ 定义 $T:V\to V$. 试求 T 的特征向量与特征值.

1.2 向量空间与几何

解析几何中熟知的坐标系允许我们去讨论几何对象, 如向量形式的直线与平面, 而这些几何概念又给我们提供了关于向量空间有益的直觉. 在考察这些几何对象的向量形式之前, 我们将简要地重复一下关于三维欧氏空间坐标对应性之构造. 通常, 成熟的读者可以略去这些材料.

我们从直线开始. 一条直线 L 与实数系 $\mathbb{R}$ 之间一个坐标对应是这样确定的, 即在 L 上任意取一个零点 O 及一个与 O 相异的单位点 Q. 则 L 上每一个点 X 就对应于一个数 x, 使 $|x|$ 等于由间隔 OQ 为单位来度量, O 至 X 的距离, 其中 x 取正值或负值依赖于 X 与 Q 在 O 之同侧或异侧. 映射 $X\mapsto x$ 就是坐标对应. 现在考虑三维欧氏空间 $\mathbb{E}^3$. 我们要建立 $\mathbb{E}^3$ 与笛卡儿向量空间 $\mathbb{R}^3$ 之间的坐标对应. 我们任意取一个零点 O 与三个单位点 Q_1,Q_2 与 Q_3 使这 4 个点不位于一张平面上, 每一个单位点 Q_i 决定一条过 O 的直线 L_i 及在这条直线上, 如上面定义的一个坐标对应. 这三条直线 L_1,L_2 与 L_3 被称为坐标轴. 现在考虑 $\mathbb{E}^3$ 中任意点 X.

过 X 且平行于 L_2 与 L_3 的平面与 L_1 在点 X_1 相交, 于是决定了一个数 x_1, 即 L_1 上 X_1 的坐标. 同法, X 决定了 L_2 上的点 X_2 及 L_3 上的点 X_3, 它们分别有坐标 x_2 与 x_3. 总之, X 决定了 $\mathbb{R}^3$ 中一个三元组

$$\boldsymbol{x} = \prec x_1, x_2, x_3 \succ,$$

所以我们定义了一个由 $\mathbb{E}^3$ 至 $\mathbb{R}^3$ 的映射 θ : $X \mapsto \boldsymbol{x}$ (见图 1.4). 我们称 θ 为由轴系统定义的坐标对应. 而上面记号中蕴涵 $\theta(Y)$ 为 $\boldsymbol{y}$ 及 $\theta(A)$ 为 $\boldsymbol{a}$ 等惯用记号. 注意 L_1 上的单位点 Q_1 有坐标三元组 $\delta^1 = \prec 1,0,0 \succ$, 及类似地有

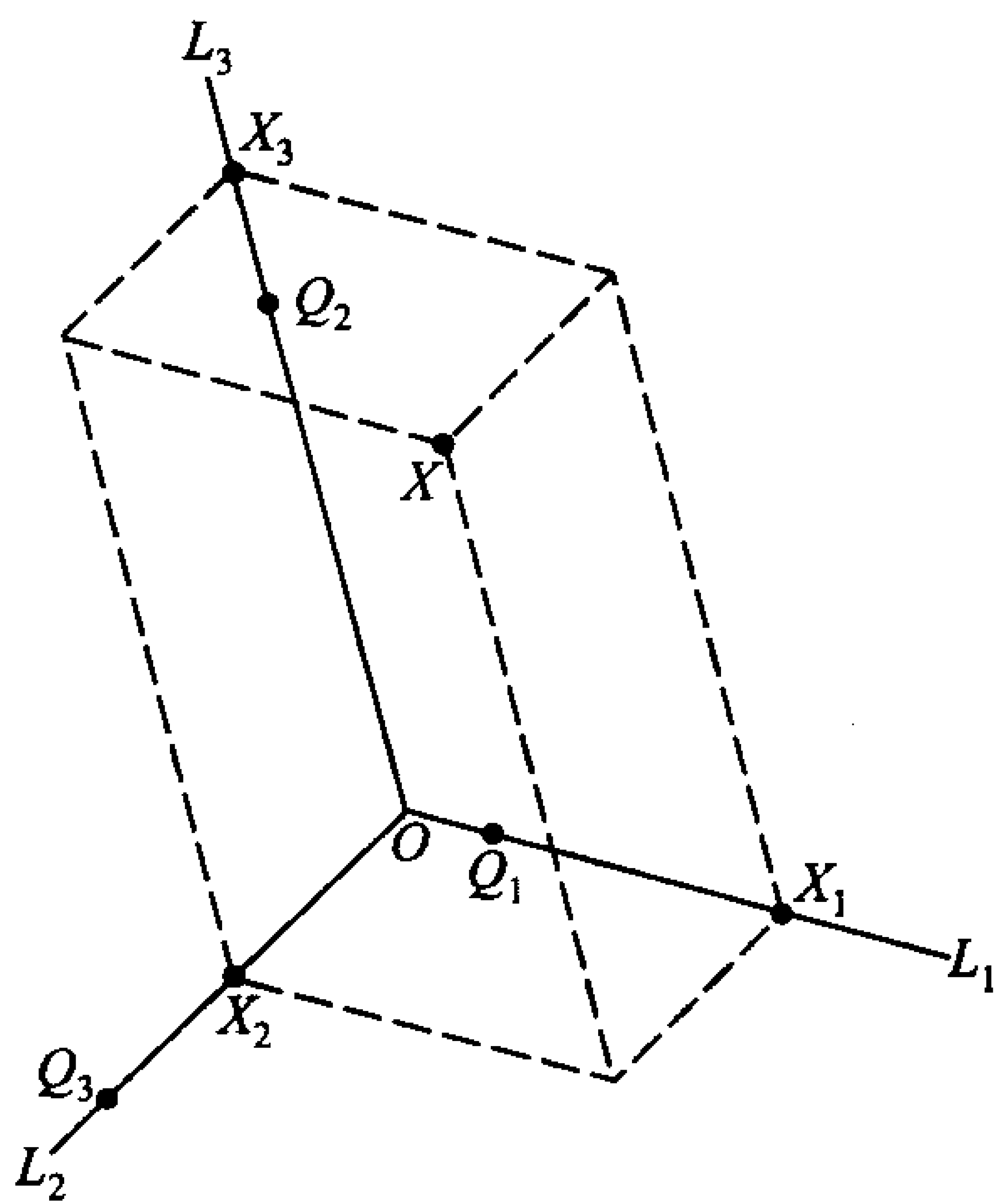


图 1.4

$$\theta(Q_2) = \delta^2 = \prec 0,1,0 \succ$$

与

$$\theta(Q_3) = \delta^3 = \prec 0,0,1 \succ .$$

在对应性可以代数地被用来处理几何问题之前已经有了一些关于坐标对应的基本事实被证明作为几何定理. 这些几何定理是很巧妙的, 它们几乎不可能根据通常中学的几何处理方法作适当的讨论. 所以我们将简单地承认它们. 它们是:

(1) θ 是由 $\mathbb{E}^3$ 至 $\mathbb{R}^3$ 的一个双射.

(2) 两个长度相等且平行的线段 AB 与 XY, 及由 A 至 B 的方向与由 X 至 Y 的方向一致当且仅当 $\boldsymbol{b} - \boldsymbol{a} = \boldsymbol{y} - \boldsymbol{x}$(在向量空间 $\mathbb{R}^3$ 中). 构筑线段间的这个关系是非常重要的. 一个有向的线段是一个几何的线段, 再加上沿着它的两个方向之一的选择. 如果我们用 AB 表示由 A 至 B 的有向线段, 若线段 AB 与 XY 的长度相等、平行且有一样的方向, 则我们定义有向线段 AB 与 XY 是等价的(并记为 $AB \sim XY$). 则 (2) 可以重述为:

$$AB \sim XY \Leftrightarrow \boldsymbol{b} - \boldsymbol{a} = \boldsymbol{y} - \boldsymbol{x}.$$

(3) 若 $X \neq O$, 则 Y 位于 $\mathbb{E}^3$ 中过 O 与 X 的直线上当且仅当有 $\mathbb{R}$ 中某 t 使 $\boldsymbol{y} = t\boldsymbol{x}$. 进而言之, 当 X 作为过 O 与 X 的直线上的单位点时, t 就是 Y 关于 X 的坐标.

(4) 若 $\mathbb{E}^3$ 中的轴系统是笛卡儿系统，即轴两两垂直并用一个相同的距离单位，则线段 OX 的长度 $|OX|$ 由 $\mathbb{R}^3$ 上所谓的欧氏范数(或模) 给出，即 $|OX| = (\sum_1^3 x_i^2)^{1/2}$. 这可以由毕达哥拉斯定理直接推出来, 然后由这公式及毕达哥拉斯定理再次用于三角形 OXY 即可推出线段 OX 与 OY 是垂直的当且仅当标量乘积$(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^3 x_i y_i$ 有值零 (见图 1.5).

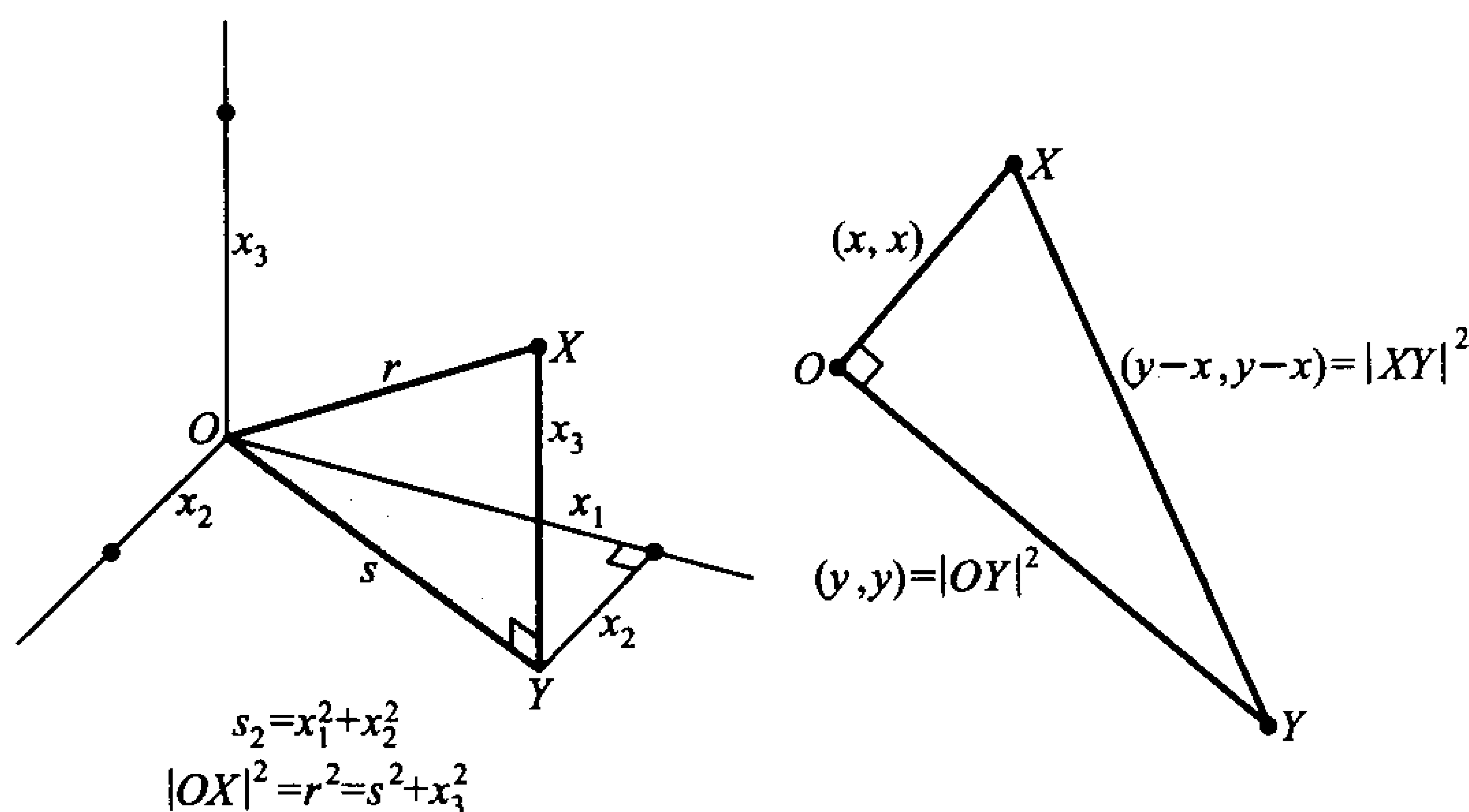


图 1.5

应用这个结果时, 注意到下述事实是有用的, 即当一个变量被固定时, 标量积 $(\boldsymbol{x},\boldsymbol{y})$ 作为另一个变量的函数是线性的, 所以

$$\begin{aligned}(c\boldsymbol{x}+d\boldsymbol{y},\boldsymbol{z}) &= \sum_1^3 (cx_i+dy_i)z_i = c\sum_1^3 x_i z_i + d\sum_1^3 y_i z_i \\ &= c(\boldsymbol{x},\boldsymbol{z}) + d(\boldsymbol{y},\boldsymbol{z}).\end{aligned}$$

对于欧氏平面 $\mathbb{E}^2$ 与笛卡儿二维空间 $\mathbb{R}^2$ 之间的坐标对应性, 当然除去现在有 $(\boldsymbol{x},\boldsymbol{y}) = \sum_1^2 x_i y_i = x_1y_1 + x_2y_2$ 之外, 完全相同的定理均成立.

由这些基本定理, 我们可以很容易得到 $\mathbb{E}^3$ 中直线与平面之方程. 首先由 (2) 与 (3) 可知, 若给定的固定点 A 与 B, 其中 $A \neq 0$, 则过 B 且平行于线段 OA 之直线包含点 X 当且仅当存在一个标量 t 使 $\boldsymbol{x}-\boldsymbol{b} = t\boldsymbol{a}$(见图 1.6). 所以这条直线的方程为

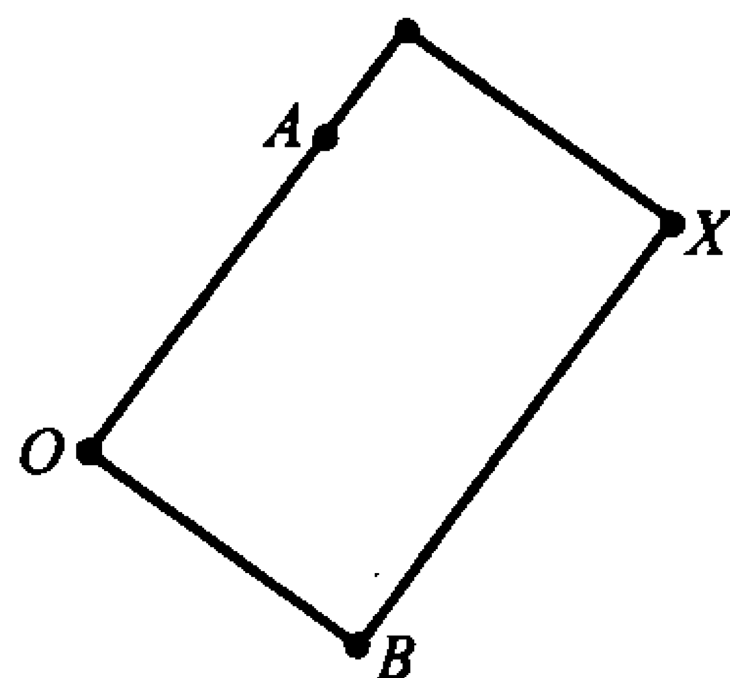


图 1.6

$$\boldsymbol{x} = t\boldsymbol{a} + \boldsymbol{b}.$$

这个向量方程等价于三个数值方程 $x_i = a_i t + b_i, i = 1, 2, 3$. 这就是习惯上所谓的*参数方程*,这是由于它们表示直线上变动点的坐标三元组 $\boldsymbol{x}$, 它是"参数"t 的函数.

其次, 我们知道过 B 且垂直于线段 OA 的方向之平面包含点 X 当且仅当 $BX \perp OA$, 所以由 (2) 与 (4) 可知这个平面包含 X 当且仅当 $(\boldsymbol{x}-\boldsymbol{b}, \boldsymbol{a}) = 0$(见图 1.7). 但由于标量乘积的第一个变量的线性性, 所以 $(\boldsymbol{x}-\boldsymbol{b}, \boldsymbol{a}) = (\boldsymbol{x}, \boldsymbol{a}) - (\boldsymbol{b}, \boldsymbol{a})$, 若我们置 $l = (\boldsymbol{b}, \boldsymbol{a})$, 则可见平面的方程为

$$(\boldsymbol{x}, \boldsymbol{a}) = l \quad \text{或} \quad \sum_{1}^{3} a_i x_i = l.$$

即一个点 X 位于过 B 且垂直于 OA 方向的平面上当且仅当 X 的坐标三元组 $\boldsymbol{x}$ 满足这个方程. 反之, 若 $\boldsymbol{a} \neq 0$, 则逆向上述步骤, 我们可以证明 $\mathbb{E}^3$ 中坐标三元组 $\boldsymbol{x}$ 满足 $(\boldsymbol{x}, \boldsymbol{a}) = l$ 的点 X 所成之集合为一张平面.

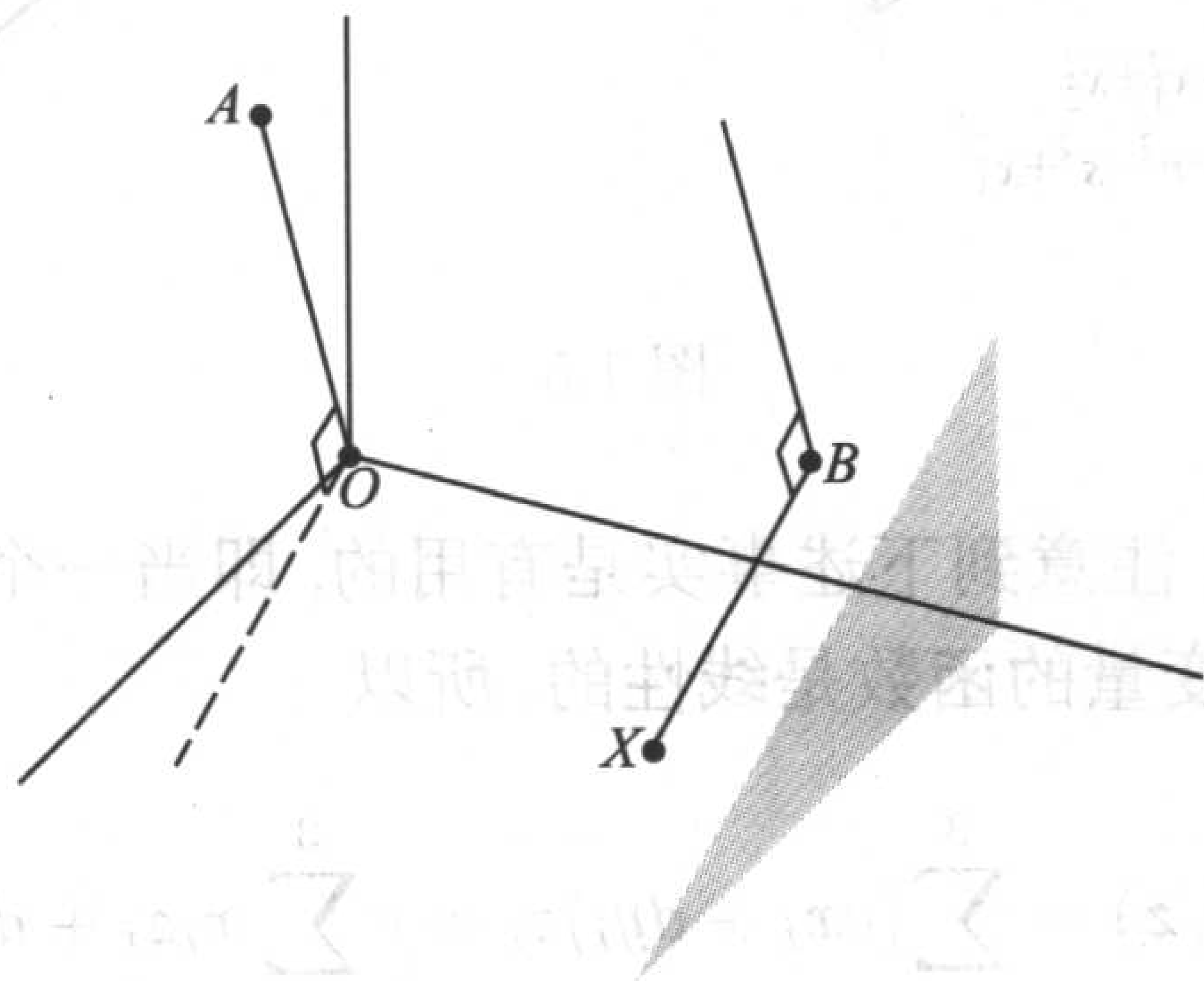


图 1.7

无论从代数角度还是几何角度, $\mathbb{R}^3$ 有自然标量乘积 $(\boldsymbol{x}, \boldsymbol{y})$ 当然都是非常重要的. 但是大多数向量空间都没有自然标量乘积, 所以我们在早先讲的向量理论中, 谨慎地避开了标量乘积 (但我们将在第五章回到这个理论). 这促使我们去寻求方程 $\sum_1^3 a_i x_i = l$ 的一个不同的表述. 在第 1 节, 我们看到 $\boldsymbol{x} \mapsto \sum_1^3 a_i x_i$ 是 $\mathbb{R}^3$ 上最一般的线性泛函 f. 所以给定 $\mathbb{E}^3$ 中任何平面 M, 存在 $\mathbb{R}^3$ 上一个非零线性泛函 f 及一个数 l 使 M 的方程为 $f(\boldsymbol{x}) = l$. 反之, 给定任意非零线性泛函 $f : \mathbb{R}^3 \to \mathbb{R}$ 及任意 $l \in \mathbb{R}, f(\boldsymbol{x}) = l$ 的轨迹是 $\mathbb{E}^3$ 中一张平面. 读者将记住由 $a_i = f(\delta^i)$, 我们从 f 得到系数三元组 $\boldsymbol{a}$, 然后就有 $f(\boldsymbol{x}) = f(\sum_1^3 x_i \delta^i) = \sum_1^3 x_i f(\delta^i) = \sum_1^3 x_i a_i$.

最后, 我们来寻求平行移动的向量形式. 在平面几何中, 当我们考虑两个平行与同向的全等图形时, 我们常常设想经 "沿着自身滑动平面" 的办法由其中的一个得到另一个; 即所有直线皆平行于它们的原始位置. 平面的一个*平行移动*的这种描述可以更精美地陈述为一条有向线段滑动至其等价者的条件. 若 X 滑动至 Y 及 O 滑动至 B, 则 OX 滑动至 BY, 所以 $OX \sim BY$, 及由 (2) 可知 $\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{b}$. 因此这种滑动的坐标形式为映射 $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{b}$.

反之, 对于 $\mathbb{R}^2$ 中任意 $\boldsymbol{b}$, 易见由 $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{b}$ 定义的平面映射为一个平行移动. 这些考虑对于欧氏空间 $\mathbb{E}^3$ 的平行移动同样成立.

从几何上看, 这是清楚的, 即在平行移动之下, 平面映射至平行平面而直线映射至平行直线, 从而我们现在可以盼望有一个简单的代数证明. 例如考虑方程为 $f(\boldsymbol{x}) = l$ 的平面 M; 我们欲问在平行移动 $\boldsymbol{x} \mapsto \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{b}$ 之下, 对 M 会发生什么? 由于 $\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{b}$, 我们得知一点 $\boldsymbol{x}$ 位于 M 上当且仅当它的位移 $\boldsymbol{y}$ 适合方程 $f(\boldsymbol{y} - \boldsymbol{b}) = l$, 或由于 f 是线性的, $\boldsymbol{y}$ 适合方程 $f(\boldsymbol{y}) = l'$, 此外 $l' = l + f(\boldsymbol{b})$. 但这是一张平面 N 的方程, 因此 M 的位移是平面 N.

很自然地将所有这些几何术语从 $\mathbb{E}^3$ 中的集合转化为 $\mathbb{R}^3$ 中对应的集合, 从而说满足 $f(\boldsymbol{x}) = l$ 的有序三元组 $\boldsymbol{x}$ 所成之集合为 $\mathbb{R}^3$ 中一个*点集*, 它构成 $\mathbb{R}^3$ 中一张平面, 并且称映射 $\boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{b}$ 为 $\mathbb{R}^3$ 中通过 $\boldsymbol{b}$ 的 (平行) 位移, 等等. 进而言之, 由于 $\mathbb{R}^3$ 为一个向量空间, 所以我们希望这些几何概念与向量概念相互作用. 例如, 经过 $\boldsymbol{b}$ 的平行移动就是简单地加上常数向量 $\boldsymbol{b}$ 的运算:$\boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{b}$. 因此若 M 为一张平面, 则经 $\boldsymbol{b}$ 平行移动 M 而得到的平面 N 正是向量集和 $M + \boldsymbol{b}$. 若 M 的方程为 $f(\boldsymbol{x}) = l$, 则平面 M 通过 O 当且仅当 $l = 0$, 在这种情况下, M 是 $\mathbb{R}^3$ 的一个*向量子空间*(f 的零空间). 易见任何平面都是一张过 O 的平面之平行移动. 类似地, 直线 $\{t\boldsymbol{a} + \boldsymbol{b} : t \in \mathbb{R}\}$ 为直线 $\{t\boldsymbol{a} : t \in \mathbb{R}\}$ 经过 $\boldsymbol{b}$ 的平行移动, 其中后一条直线是一个子空间, 即一个向量 $\boldsymbol{a}$ 的线性生成. 因此 *$\mathbb{R}^3$ 中平面与直线都是子空间的平行移动*.

所有这些概念均可以用一种完备与满意的方式转至任意实向量空间, 并有附加的维数变化. $\mathbb{R}^3$ 中过 O 的一张平面是一个向量空间, 它在严格的代数意义之下是二维的. 我们将在下一章讨论它. 类似地, 一条直线是一维的. 在 $\mathbb{R}^3$ 中, 除了过 O 的平面与直线之外, 没有真子空间. 但是在维数 > 3 的一个向量空间 V 中, 有维数从 1 至 $n-1$ 的所有真子空间. 因此我们将用术语 "平面" 含糊地表示不论其维数的一个子空间的任意平行移动. 更正规地, 向量子空间的平行移动称为*仿射子空间*.

我们将看到如果 V 是一个维数 n 的有限维空间, 则一个非零线性泛函 f 的

零空间总是 $(n-1)$ 维的. 从而除 $n=3$ 之外, 它不可能是一个类欧氏二维平面. 我们用术语*超平面*来表示这样一个零空间或它的一个平行移动. 因此, 一般说来, 一个超平面是一个适合方程 $f(\boldsymbol{x})=l$ 之 $\boldsymbol{x}$ 的集合, 此处 f 为一个非零线性泛函. 它是一个真仿射子空间 (平面), 它在下面意义之下是极大的, 即仅有的真正包含它的仿射子空间为整个 V. 在 $\mathbb{R}^3$ 中, 超平面为普通的几何平面, 而在 $\mathbb{R}^2$ 中, 超平面则为直线！

习题

2.1 假定定理 $AB \sim XY \Leftrightarrow \boldsymbol{b}-\boldsymbol{a}=\boldsymbol{y}-\boldsymbol{x}$ 成立. 试证如第 1 节开始讨论时所定义的 $\overrightarrow{OC}$ 是 $\overrightarrow{OA}$ 与 $\overrightarrow{OB}$ 之和当且仅当 $\boldsymbol{c}=\boldsymbol{b}+\boldsymbol{a}$. 又考虑我们所假定的几何定理 (3), 试证由 $\mathbb{R}^3$ 至几何向量所成空间的映射 $\boldsymbol{x}\mapsto\overrightarrow{OX}$ 是线性的, 从而是一个同构.

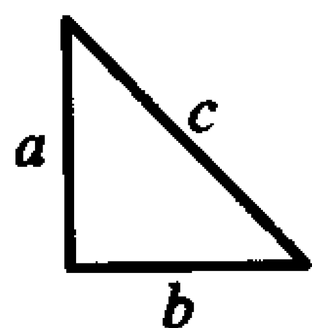
2.2 命 L 为笛卡儿平面 $\mathbb{R}^2$ 上有方程 $x_2=3x_1$ 之直线. 试用一个适当的序对 $\boldsymbol{a}$ 将 L 表示为形如 $\boldsymbol{x}=t\boldsymbol{a}$ 的参数形式.

2.3 命 V 为任意向量空间及 α 与 β 为相异的向量. 试证过 α 与 β 的直线有参数方程

$$\xi=t\beta+(1-t)\alpha,\quad t\in\mathbb{R}.$$

再证明由 α 至 β 的间隔是上面映射中, [0,1] 的像.

2.4 按照毕达哥拉斯定理, 一个三角形其边长为 a,b 与 c,"对着 c" 的顶角为直角当且仅当 $c^2=a^2+b^2$.



试由此证明在 $\mathbb{E}^3$ 的笛卡儿坐标系中, 线段 OX 的长度 $|OX|$ 等于

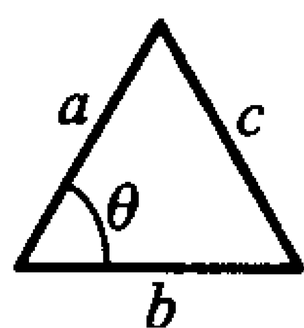
$$|OX|=\sum_1^3 x_i^2,$$

此处 $\boldsymbol{x}=\prec x_1,x_2,x_3\succ$ 为点 X 的坐标三元组. 其次试用我们的几何定理 (2) 证明

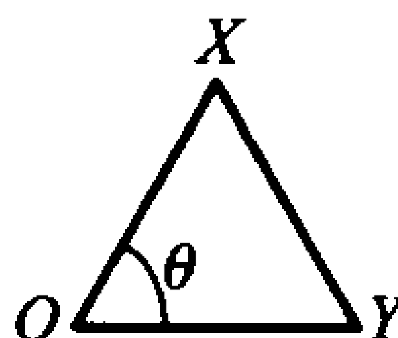
$$OX\perp OY\Leftrightarrow(\boldsymbol{x},\boldsymbol{y})=0,\quad \text{其中}(\boldsymbol{x},\boldsymbol{y})=\sum_1^3 x_iy_i.$$

(利用 $(\boldsymbol{x},\boldsymbol{y})$ 的双线性性来展开 $|X-Y|^2$) .

2.5 更一般些, 对于下图所示的任意三角形,



余弦定律是说 $c^2 = a^2 + b^2 - 2ab\cos\theta$. 将这一定律用于图形



证

$$(\boldsymbol{x}, \boldsymbol{y}) = 2|\boldsymbol{x}||\boldsymbol{y}|\cos\theta,$$

此处 $(\boldsymbol{x},\boldsymbol{y})$ 表示标量乘积 $\sum_1^3 x_i y_i$, $|\boldsymbol{x}| = (\boldsymbol{x}, \boldsymbol{x})^{1/2} = |OX|$, 等等.

2.6 给定一个非零线性泛函 $f : \mathbb{R}^3 \to \mathbb{R}$ 及 $k \in \mathbb{R}$, 试证 $\mathbb{E}^3$ 中满足 $f(\boldsymbol{x}) = k$ 的点 X 之集合是一张平面 [提示: 在 $\mathbb{R}^3$ 中找出一个 $\boldsymbol{b}$ 使 $f(\boldsymbol{b}) = k$, 及将方程 $f(\boldsymbol{x}) = k$ 化为形式 $(\boldsymbol{x} - \boldsymbol{b}, \boldsymbol{a}) = 0$ 等等].

2.7 对于任意 $\mathbb{R}^3$ 中的 $\boldsymbol{b}$, 试证由 $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{b}$ 定义的将 $\mathbb{E}^3$ 映为自身的映射 $X \mapsto Y$ 是一个平行移动. 即证若 $X \mapsto Y$, 及 $Z \mapsto W$, 则 $XZ \sim YW$.

2.8 命 M 为 $\mathbb{R}^3$ 中适合方程 $3x_1 - x_2 + x_3 = 2$ 的点所成之集合. 试找出三元组 $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 使 M 为过 $\boldsymbol{b}$ 且与 $\boldsymbol{a}$ 的方向垂直的平面. 平面 $P = M + \prec 1, 2, 1 \succ$ 的方程是什么?

2.9 继续上面习题, 什么是三元组 $\boldsymbol{b}$ 满足的条件使 $N = M + \boldsymbol{b}$ 经过原点?N 的方程是什么?

2.10 试证若 $\mathbb{R}^3$ 中的平面 M 有方程 $f(\boldsymbol{x}) = l$, 则 M 是线性泛函 f 的零空间的一个平移. 再证 N 的任何两个平移 M 与 P 或者恒等, 或者不相交. 什么是有序三元组 $\boldsymbol{b}$ 需满足的条件使 $M + \boldsymbol{b} = M$?

2.11 将上面习题推广至 $\mathbb{R}^n$ 中的超平面.

2.12 命 N 为 $\mathbb{R}^3$ 中满足 $f(\boldsymbol{x}) = 0$ 的点 $\boldsymbol{x}$ 所成之子空间 (过原点的平面). 命 M 与 P 为由 N 经平行移动得到的两张平面. 试证 $Q = M + P$ 为第三张这样的平面. 若 M 与 P 有方程 $f(\boldsymbol{x}) = l_1$ 与 $f(\boldsymbol{x}) = l_2$, 试求 Q 的方程.

2.13 若 M 为 $\mathbb{R}^3$ 中有方程 $f(\boldsymbol{x}) = l$ 之平面及 r 为任意非零数. 试证集合乘积 rM 是平行于 M 的一张平面.

2.14 由上面两个习题, 试讨论我们怎样才可以将具有方程 $f(\boldsymbol{x}) = 0$ 的平面 N 的所有平行移动所成之集合作为构成一个新的向量空间.

2.15 命 L 为 $\mathbb{R}^3$ 中具有参数方程 $\boldsymbol{x} = t\boldsymbol{a}$ 的子空间 (过原点之直线). 试用上面 3 个习题之精神讨论 L 的所有平行移动所成之集合.

2.16 作为 "是" 几何向量 $\overrightarrow{AB}$ 的最好的事物是所有满足 $XY \sim AB$ 的有向线段 XY 的等价类. 在性质 (1)~(4) 中, 假定你所需要者都成立, 试证这是所有有向线段的集合上的一个等价关系 (0.12 节).

2.17 假定几何向量 $\overrightarrow{AB}$ 已于上面习题所定义. 试证, 严格地讲, 这实际上是平面 (或空间) 到自身的映射, 此处自身是我们所称的过 $\overrightarrow{AB}$ 的平行移动. 再证 $\overrightarrow{AB} + \overrightarrow{CD}$ 是两个平移的合成.

1.3 积空间与 $\mathrm{Hom}(V, W)$

积空间 若 W 是一个向量空间及 A 是一个任意集合, 则 A 上所有 W 值函数所成的集合 $V = W^A$ 正好是与 $\mathbb{R}^A$ 由同一途径得来的一个向量空间. 加法是函数的自然加法, $(f+g)(a) = f(a) + g(a)$, 及类似地, 对于每一个函数 f 与标量 x 有 $(xf)(a) = x(f(a))$. 恰如以前的同样理由可以推出定律 A1 到 S4 成立. 为了广泛性, 让我们来验证加法结合律. 方程 $f + (g+h)$ 的意思是, $(f+(g+h))(a) = ((f+g)+h)(a)$ 对于所有 $a \in A$ 成立. 事实上,

$$\begin{aligned}(f+(g+h))(a) &= f(a) + (g+h)(a)\\ &= f(a) + (g(a)+h(a)) = (f(a)+g(a)) + h(a)\\ &= (f+g)(a) + h(a) = ((f+g)+h)(a),\end{aligned}$$

此处由 W 的结合律可知这 5 个等式链的中间等式成立, 而其他 4 个等式则是加法定义的应用. 由于加法结合律在 W 中成立, 所以它在 W^A 中亦成立. 完全同样的方法可知其他定律成立. 如前, 我们命 π_i 为 i 处的求值, 所以 $\pi_i(f) = f(i)$. 但是现在 π_i 是向量值而不是标量值. 这是由于它是一个由 V 到 W 的映射, 所以我们称它为第 i 个坐标投影,而不是第 i 个坐标泛函. 这些映射仍然都是线性的. W^A 上的自然向量运算是被要求所有投影 π_i 都是线性的而惟一地定义的. 我们称 $f(j) = \pi_j(f)$ 为向量 f 的第 j 个坐标. 在此笛卡儿 n 维空间的类似是 W 中所有的向量的 n 元组 $\boldsymbol{\alpha} = \prec \alpha_1, \cdots, \alpha_n \succ$ 所成的集合 $W^{\overline{n}}$; 它们亦被记为 W^n. 显然, α_j 是 n 元组 $\boldsymbol{\alpha}$ 的第 j 个坐标.

我们为什么必须如上面这样用同样的空间 W 于每一个指标是没有理由的. 事实上, 若 $W_1, \cdots, W_n$ 是任意 n 个向量空间, 则所有 n 元组 $\boldsymbol{\alpha} = \prec \alpha_1, \cdots, \alpha_n \succ$

所成的集合, 此处当 $j = 1, \cdots, n$ 时, $\alpha_j \in W_j$, 在运算的同样定义及同样的理由下, 是一个向量空间, 即笛卡儿乘积 $W = W_1 \times W_2 \times \cdots \times W_n$ 亦是一个向量值函数的向量空间. 这种有限乘积对我们将是非常重要的. 当然, $\mathbb{R}^n$ 就是乘积 $\prod_1^n W_i$, 其中每一个 $W_i = \mathbb{R}$; 但是 $\mathbb{R}^n$ 还可以被当作 $\mathbb{R}^m \times \mathbb{R}^{n-m}$, 或更一般些, $\prod_1^p W_i$, 其中 $W_i = \mathbb{R}^{m_i}$ 及 $\sum_1^p m_i = n$. 无论如何, 有限积空间最重要的应用来自这样一个事实, 即一个向量空间 V 的某些现象的研究可以由一种自然的方式导致 V 的子空间的一个集合 $\{V_i\}_1^n$ 使 V 同构于乘积 $\prod_1^n V_i$. 当我们将 V 当作积空间 $\prod_1^n V_i$ 时, V 有的额外结构就可以被用来研究问题中的现象. 这就是直和理论, 我们将在第 5 节研究它.

在教程的后面, 我们将需要考虑向量空间的一个一般笛卡儿乘积. 我们提醒读者, 若 $\{W_i : i \in I\}$ 为任意向量空间的有指标集合, 则这些向量空间的笛卡儿乘积 $\prod_{i\in I} W_i$ 被定义为以 I 为定义域且满足对于所有 $i \in I$, $f(i) \in W_i$ 的所有函数 f 所成之集合 (见 0.8 节).

下面是一个简单的具体例子, 需记在心上, 命 S 为 $\mathbb{R}^3$ 中通常的单位球, $S = \{\boldsymbol{x} : \sum_1^3 x_i^2 = 1\}$, 及对于每一个 S 上的点 $\boldsymbol{x}$, 命 $W_{\boldsymbol{x}}$ 为 $\mathbb{R}^3$ 中在 $\boldsymbol{x}$ 与 S相切的子空间. 我们在此关于子空间的意思是 (过 O 的平面) 平行于在 $\boldsymbol{x}$ 处 S 的切平面. 所以平移 $W_{\boldsymbol{x}} + \boldsymbol{x}$ 为切平面 (见图 1.8). 积空间 $W = \prod_{\boldsymbol{x}\in S} W_{\boldsymbol{x}}$ 中一个函数 f 是对 S 上每一点 $\boldsymbol{x}$, 给出 $W_{\boldsymbol{x}}$ 中一个向量元函数, 即一个平行于 S 在 $\boldsymbol{x}$ 的切平面之向量. 这样一个函数被称为 S 上的一个向量场.因此乘积集合 W 是 S 上所有向量场所成的集合. 如下面定理所述, W 自身就是一个向量空间.

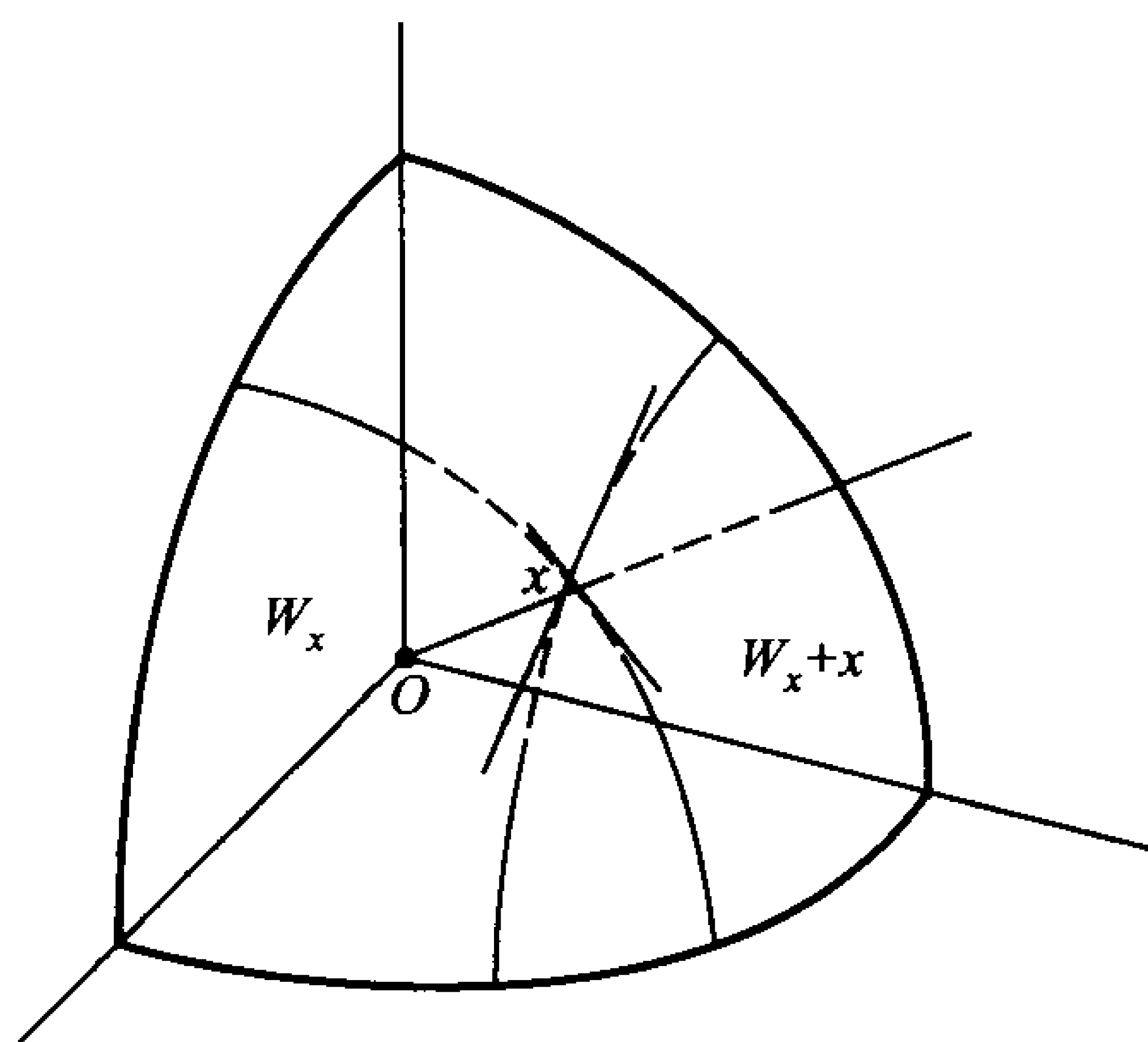


图 1.8

当然, 在 $W = \prod_{i \in S} W_i$ 上的第 j 个坐标投影是在 j 的求值, $\pi_j(f) = f(j)$, 及 W 上的自然向量运算是坐标投影全是线性的要求所惟一地定义的. 因此 $f + g$ 必须是 W 的那个元素, 它在 j 处之值 $\pi_j(f+g)$ 是 $\pi_j(f) + \pi_j(g) = f(j) + g(j)$, 此处 j 为任意 $j \in I$. 关于标量乘积有类似定义.

定理 3.1 一个向量空间集合的笛卡儿乘积正好由一种方式可以进入一个向量空间, 使坐标投影都是线性的.

证明 由上面惟一确定的向量运算, 关于 A1~S4 的证明, 我们可以将过去的证明逐字地重述. 它们不需要函数相加后, 其所有值均在同一空间中, 但仅仅要在给定的一个定义域元素 i 之值位于同样空间中. □

Hom(V, W) 线性变换有简单而重要的性质, 即两个线性变换之和是线性的, 及二个线性变换之合成亦是线性的. 这些不精确的语句本质上就是本节之论题, 尽管它们还需要放上定义域与值域的条件. 它们的证明则是简单的形式代数推导, 但讨论的对象在概念复杂性上将有所增加.

若 W 是一个向量空间及 A 为任意集合, 如同 $\mathbb{R}^A$ 一样, 我们知道所有映射 $f: A \to W$ 所成的空间 W^A 是一个函数的向量空间 (现在是向量值的). 若 A 自身是一个向量空间 V, 则我们自然地挑出 W^V 中包含所有线性映射的子集作特殊的研究. 我们将这个子集合记作 $\operatorname{Hom}(V, W)$. 下面一些初等定理概括了它的基本代数性质.

定理 3.2 $\operatorname{Hom}(V, W)$ 是 W^V 的向量子空间.

证明 这个定理是一个容易的形式化结果。若 S 与 T 属于 $\operatorname{Hom}(V, W)$, 则

$$
\begin{aligned}
(S+T)(x\alpha + y\beta) &= S(x\alpha + y\beta) + T(x\alpha + y\beta) \\
&= xS(\alpha) + yS(\beta) + xT(\alpha) + yT(\beta) \\
&= x(S+T)(\alpha) + y(S+T)(\beta),
\end{aligned}
$$

所以 $S + T$ 是线性的及 $\operatorname{Hom}(V, W)$ 在加法之下是闭的. 读者应该被认为他知道辨明上面连续恒等式中的每一步. 在标量乘积下, $\operatorname{Hom}(V, W)$ 的闭包可以类似地得到. 由于 $\operatorname{Hom}(V, W)$ 包含零变换, 所以它是非空的, 从而是一个子空间. □

定理 3.3 线性映射的合成是线性的: 若 $T \in \operatorname{Hom}(V, W)$ 及 $S \in \operatorname{Hom}(W, X)$, 则 $S \circ T \in \operatorname{Hom}(V, X)$. 进而言之, 在定义域与值域的显然假设下, 合成适合

加法分配律:

$$(S_1+S_2)\circ T=S_1\circ T+S_2\circ T \quad 与 \quad S\circ(T_1+T_2)=S\circ T_1+S\circ T_2.$$

最后, 合成可与标量乘积交换:

$$c(S\circ T)=(cS)\circ T=S\circ(cT).$$

证明 我们有

$$\begin{aligned}S\circ T(x\alpha+y\beta)&=S(T(x\alpha+y\beta))=S(xT(\alpha)+yT(\beta))\\&=xS(T(\alpha))+yS(T(\beta))=x(S\circ T)(\alpha)+y(S\circ T)(\beta),\end{aligned}$$

所以 $S\circ T$ 是线性的. 两条分配律的证明将留给读者. □

推论 3.1 若 $T\in \mathrm{Hom}(V,W)$ 为固定的, 则由 T 在右方之合成是一个从向量空间 Hom (W, X) 至向量空间 $\mathrm{Hom}(V,X)$ 的线性变换. 若 T 是一个同构, 则这个变换就是一个同构.

证明 定理中所述合成的代数性质可以综合如下:

$$(c_1S_1+c_2S_2)\circ T=c_1(S_1\circ T)+c_2(S_2\circ T),$$

$$S\circ(c_1T_1+c_2T_2)=c_1(S\circ T_1)+c_2(S\circ T_2).$$

另一个方程正好是说, 一个固定的 T 在右方的合成是一个线性变换 (若方程仍然没有看对, 则记 $S\circ T=\mathcal{T}(S)$). 若 T 是一个同构, 则 T^{-1} 之合成 "取消" 了由 T 之合成, 所以是它的逆元.

第二个方程推出一个由一个固定 S 从左方合成的类似的推论. □

定理 3.4 若 W 为一个积向量空间, $W=\prod_i W_i$, 则由一个向量空间 V 至 W 的映射 T 为线性的当且仅当对于每一个坐标投影 π_i, $\pi_i\circ T$ 是线性的.

证明 若 T 是线性的, 则由上面定理可知 $\pi_i\circ T$ 是线性的. 现在假定反过来, 即所有映射 $\pi_i\circ T$ 都是线性的. 则

$$\begin{aligned}\pi_i(T(x\alpha+y\beta))&=\pi_i\circ T(x\alpha+y\beta)=x(\pi_i\circ T)(\alpha)+y(\pi_i\circ T)(\beta)\\&=x\pi_i(T(\alpha))+y\pi_i(T(\beta))=\pi_i(xT(\alpha)+yT(\beta)).\end{aligned}$$

但是若 $\pi_i(f)=\pi_i(g)$ 对所有 i 成立, 则 $f=g$, 所以 $T(x\alpha+y\beta)=xT(\alpha)+yT(\beta)$, 从而 T 是线性的. □

若 T 是一个由 $\mathbb{R}^n$ 至 W 的线性映射, 其骨架为 $\{\beta_j\}_1^n$, 则 $\pi_i \circ T$ 有骨架 $\{\pi_i(\beta_j)\}_{j=1}^n$. 若 W 是 $\mathbb{R}^m$, 则 π_i 为第 i 个坐标泛函 $\boldsymbol{y} \mapsto y_i$ 及 β_j 为 T 的矩阵 $\boldsymbol{t} = \{t_{ij}\}$ 的第 j 列. 因此 $\pi_i(\beta_j) = t_{ij}$, 及 $\pi_i \circ T$ 为线性泛函, 其骨架为 T 的矩阵的第 i 行.

在集中围绕定理 1.3 的讨论中, 我们将向量方程 $\boldsymbol{y} = T(\boldsymbol{x})$ 换成等价的 m 个标量方程 $y_i = \sum_{j=1}^n t_{ij}x_j$ 的集合, 在此我们可以由向量方程中, 读出第 i 个坐标而得到. 但是在 "读出" 第 i 个坐标时, 我们是应用坐标映射 π_i, 或用更代数的术语, 我们将线性映射 T 换成线性映射之集合 $\{\pi_i \circ T\}$, 由上面的定理可知它等价于 T.

我们现在特别考虑空间 $\mathrm{Hom}(V, V)$, 我们也可以将它记为 '$\mathrm{Hom}\, V$'. 除了作为一个向量空间之外, 它在合成之下也是闭的, 我们将它考虑为一个乘法运算. 由于函数的合成总是可结合的 (见 0.9 节), 所以我们有关于乘法的定律:

$$A \circ (B \circ C) = (A \circ B) \circ C,$$

$$A \circ (B + C) = (A \circ B) + (A \circ C),$$

$$(A + B) \circ C = (A \circ C) + (B \circ C),$$

$$k(A \circ B) = (kA) \circ B + A \circ (kB).$$

任何向量空间, 除了有向量运算之外, 还有如上方式的与向量运算相关之乘法运算, 则称为一个代数. 因此有

定理 3.5　$\mathrm{Hom}(V)$ 是一个代数.

我们在较早时注意到某些实值函数空间亦为代数. 例如 $\mathbb{R}^A$ 与 $\mathfrak{C}([0,1])$. 在这些情况下, 乘法是可交换的, 但在 $\mathrm{Hom}(V)$ 的情况下, 除了 V 是一个平凡空间 $(V = \{0\})$ 或 V 同构于 $\mathbb{R}$, 乘法运算都是不可交换的. 当我们详细考察有限维理论时, 我们将在以后检验这一点.

乘积投影与单射　除了坐标投影之外, 还有第二类简单的线性映射, 它对于掌握一个笛卡儿乘积空间 $W = \prod_{k \in K} W_k$ 有基本重要性. 对于每一个 j, 有映射 θ_j; 取一个向量 $\alpha \in W_j$ 至乘积空间中之函数, 它在指标 j 处取值 α, 否则取值 0, 例如, 关于 $W_1 \times W_2 \times W_3$, θ_2 是由 W_2 至 W 的映射 $\alpha \mapsto \prec 0, \alpha, 0 \succ$. 或者若我们将 $\mathbb{R}^3$ 看成 $\mathbb{R} \times \mathbb{R}^2$, 则 θ_2 就是映射 $\prec x_2, x_3 \succ \mapsto \prec 0, \prec x_2, x_3 \succ \succ = \prec 0, x_2, x_3 \succ$. 我们称 θ_j 为 W_j 至 $\prod_k W_k$ 的单射映射. θ_j 的线性性可能是显然的. 映射 π_j 与

θ_j 显然是相关联的. 下面的投影 - 单射恒等关系陈述了它们的准确关系. 若 I_j 为 W_j 上的恒等变换, 则

$$\pi_j \circ \theta_j = I_j \quad 及 \quad \pi_j \circ \theta_i = 0 (i \neq j).$$

若 K 为有限的 及 I 为乘积空间 W 上的单位元 (或恒等变换), 则

$$\sum_{k \in K} \theta_k \circ \pi_k = I.$$

在 $\prod_{i=1}^{3} W_i$ 的情况下, 我们有 $\theta_2 \circ \pi_2(\prec \alpha_1, \alpha_2, \alpha_3 \succ) = \prec 0, \alpha_2, 0 \succ$, 及恒等变换简单地说就是 $\prec \alpha_1, 0, 0 \succ + \prec 0, \alpha_2, 0 \succ + \prec 0, 0, \alpha_3 \succ = \prec \alpha_1, \alpha_2, \alpha_3 \succ$ 对于所有 $\alpha_1, \alpha_2, \alpha_3$ 成立. 这些恒等关系对读者来说可能是清楚的, 我们将正式的证明留作习题.

坐标投影 π_j 对于任何乘积空间的研究都是有用的. 但是因为上面恒等关系的限制, 单射映射 θ_j 主要的兴趣在于有限乘积的情况. 它们放在一起使我们能够分解与聚合线性映射, 此处映射的定义域与值域是有限乘积空间.

作为一个简单的例子, 考虑 $\mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^2)$ 中的 T, 其矩阵为

$$\begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 4 \end{bmatrix}.$$

则 $\pi_1 \circ T$ 是线性泛函, 其骨架 $\prec 2, -1, 1 \succ$ 是 T 的矩阵中的第 1 行, 而且我们能够看出它的方程形式表示式 $y_1 = 2x_1 - x_2 + x_3$ 是从向量方程 $\boldsymbol{y} = T(\boldsymbol{x})$ 中"读出第一行" 而得到的. 因此我们 "分解"T 为两个线性泛函 $l_i = \pi_i \circ T$. 则模糊地说, 我们有聚合 $T = \prec l_1, l_2 \succ$; 更确切地, 对于所有 $\boldsymbol{x}$, $T(\boldsymbol{x}) = \prec 2x_1 - x_2 + x_3, x_1 + x_2 + 4x_3 \succ = \prec l_1(\boldsymbol{x}), l_2(\boldsymbol{x}) \succ$. 但是我们要这个聚合表示为线性映射 θ_1 与 θ_2 的作用。我们有

$$\prec l_1(\boldsymbol{x}), l_2(\boldsymbol{x}) \succ = \theta_1(l_1(\boldsymbol{x})) + \theta_2(l_2(\boldsymbol{x})) = (\theta_1 \circ \pi_1 + \theta_2 \circ \pi_2)(T(\boldsymbol{x})) = T(\boldsymbol{x}),$$

这证明了 T 的分解与聚合是恒等映射 $\sum \theta_i \circ \pi_i = I$ 的一个表达式. 一般言之, 若 $T \in \mathrm{Hom}(V, W)$ 及 $W = \prod_i W_i$, 则对于每一个 i, $T_i = \pi_i \circ T$ 在 $\mathrm{Hom}(V, W_i)$ 之中, 及由于对于每一个 α, $T_i(\alpha)$ 是 $T(\alpha)$ 的第 i 个坐标, 所以 T_i 可以被当成是 "T 进入 W_i 的部分". 则我们可以由 $T = \sum \theta_i \circ T_i$ 再将 T_i 聚合起来形成 T, 这里 $\sum \theta_i \circ T_i = (\sum \theta_i \circ \pi_i) \circ T = I \circ T = T$. 进而言之, 在一个共同的定义域上, T_i 的任意有限集合皆可以用这个方法放在一起去造成一个 T. 例如, 我们可

以在一个共同定义域 V 上聚合线性变换 m 元组 $\{T_i\}_1^m$ 去构成一个单个的 m 元值线性变换 T. 给予 V 中 α, 我们简单地定义 $T(\alpha)$ 为 m 元组, 它的第 i 个坐标为 $T_i(\alpha), i=1,\cdots,m$, 然后再验证 T 是线性的. 因此, 不作计算, 我们由这种聚合原则看到 $T:\boldsymbol{x}\mapsto \prec 2x_1-x_2+x_3, x_1+x_2+4x_3\succ$ 是一个由 $\mathbb{R}^3$ 至 $\mathbb{R}^2$ 的线性映射, 这是由于我们已经由聚合两个线性泛函 $l_1(\boldsymbol{x})=2x_1-x_2+x_3$ 与 $l_2(\boldsymbol{x})=x_1+x_2+4x_3$ 组成了 T, 即形成一个有序的双值映射. 这个非常直观的程序有一个同样简单的形式上的证明. 我们将严格的讨论放在下面定理中.

定理 3.6 若对于一个有限指标集合 I 中的每一个 i, T_i 都在 $\mathrm{Hom}(V, W_i)$ 之中, 及若 W 是乘积空间 $\prod_{i\in I} W_i$, 则在 $\mathrm{Hom}(V, W)$ 中有惟一确定的 T 使 $T_i=\pi_i\circ T$ 对所有 $i\in I$ 成立.

证明 若存在 T 使对于每一个 i, $T_i=\pi_i\circ T$, 则 $T=I_W\circ T=(\sum\theta_i\circ\pi_i)\circ T=\sum\theta_i\circ(\pi_i\circ T)=\sum\theta_i\circ T_i$. 因此, T 被惟一地确定为 $\sum\theta_i\circ T_i$. 进而言之, 这个 T 是有所要求之性质, 这是由于

$$\pi_j\circ T=\pi_j\circ(\textstyle\sum\theta_i\circ T_i)=\sum_i(\pi_j\circ\theta_i)\circ T_i=I_j\circ T_j=T_j. \qquad \square$$

同法, 我们可以分解一个其定义域为一个乘积空间 $V=\prod_{j=1}^n V_i$ 的线性映射 T 为诸定义域为 V_j 的映射 $T_i=T\circ\theta_j$, 然后用恒等关系 $T=\sum_{j=1}^n T_j\circ\pi_j$ 聚合这些映射来形成 T(用心算检验一下!). 进而言之, 一个映至共同上域空间的有限映射集合可以放在一起来形成一个在定义域乘积上的单个映射. 因此, 一个映射至 W 的 n 元组映射 $\{T_i\}_1^n$ 定义了一个映至 W 的单个映射 T, 此处由方程 $T(\prec\alpha_1,\cdots,\alpha_n\succ)=\sum_1^n T_i(\alpha_i)$ 或 $T=\sum_1^n T_i\circ\pi_i$ 可知 T 的定义域为各 T_i 的定义域之乘积. 例如, 如果 $T_1:\mathbb{R}\mapsto\mathbb{R}^2$ 为映射 $t\mapsto t\prec 2,1\succ=\prec 2t,t\succ$, 及类似地, T_2 与 T_3 为映射 $t\mapsto t\prec -1,1\succ$ 与 $t\mapsto t\prec 1,4\succ$, 则 $T=\sum_1^3 T_i\circ\pi_i$ 为由 $\mathbb{R}^3$ 至 $\mathbb{R}^2$ 的映射, 其矩阵为

$$\begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 4 \end{bmatrix}.$$

这里亦有一个简单的形式论证, 我们将要求读者写出下面定理的证明.

定理 3.7 若对于一个有限指标集合 J 的每一个 j, T_j 皆在 $\mathrm{Hom}(V_j, W)$ 之中, 及若 $V=\prod_{j\in J} V_j$, 则存在 $\mathrm{Hom}(V, W)$ 中一个惟一的 T 使对于 J 中每一个 j 皆有 $T\circ\theta_j=T_j$.

最后, 我们应该提到定理 3.6 对于所有乘积空间皆成立. 包括有限与无限的乘积空间, 这陈述了特征化乘积空间的一个性质. 我们将在习题中研究这种情况. 定理 3.6 的一般情况的证明是按没有单射 θ_j 而得到的; 实际上, 这是定理 3.4 的应用.

读者可能感到我们过度形式地在应用投影 π_i 与单射 θ_i 去给出程序的代数表述, 这程序是易于被直接看出来的, 例如读出一个向量方程的标量 "分量". 但是, 映射

$$\boldsymbol{x} \mapsto x_i \quad 与 \quad x_i \mapsto \prec 0, \cdots, 0, x_i, 0, \cdots, 0 \succ$$

显然是基本的设计, 及使它们的关系显式化. 在以后更为复杂的情况下，当我们要去掌握它们的发生时，将会有帮助的。

习题

3.1 试证 $\mathbb{R}^m \times \mathbb{R}^n$ 同构于 $\mathbb{R}^{n+m}$.

3.2 更一般些，试证若 $\sum_1^m n_i = n$, 则 $\prod_{i=1}^m \mathbb{R}^{n_i}$ 同构于 $\mathbb{R}^n$.

3.3 试证若 $\{B, C\}$ 是 A 的一个分拆，则 $\mathbb{R}^A$ 与 $\mathbb{R}^B \times \mathbb{R}^C$ 是同构的。

3.4 试将上面习题推广为 $\{A_i\}_1^r$ 分拆 A 的情况.

3.5 试证由一个向量空间 V 至一个向量空间 W 的一个映射是线性的当且仅当 T(的图) 是 $V \times W$ 的子空间.

3.6 命 S 与 T 为由 V 至 W 的非零线性映射. 映射 $S+T$ 的定义与作为 $V \times W$ 的子空间 S 与 T(的图) 的集和是不相同的. 试证除 $S = T$ 之外, S 与 T(的图) 的集和不能是一个图.

3.7 试给出定理 3.2 中计算的每一步的论证.

3.8 试证定理 3.3 中的分配律.

3.9 命 $D : \mathfrak{C}^1([a, b]) \to \mathfrak{C}([a, b])$ 为微分, 及命 $S : \mathfrak{C}([a, b]) \to \mathbb{R}$ 为定积分映射 $f \mapsto \int_a^b f$. 试计算合成 $S \circ D$.

3.10 我们知道 $\mathbb{R}^2$ 上的一般线性泛函 F 是由 $\mathbb{R}^2$ 中数对 $\boldsymbol{a}$ 决定的映射 $\boldsymbol{x} \mapsto a_1x_1 + a_2x_2$, 及 $\mathrm{Hom}(\mathbb{R}^2)$ 中的一般线性映射 T 是由矩阵

$$\boldsymbol{t} = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}$$

决定的. 则 $F \circ T$ 为另一个线性泛函, 因此有形式 $\boldsymbol{x} \mapsto b_1x_1 + b_2x_2$, 其中某 $\boldsymbol{b} \in \mathbb{R}^2$. 试由 $\boldsymbol{t}$ 与 $\boldsymbol{a}$ 来计算 $\boldsymbol{b}$. 你的计算应当向你显示 $\boldsymbol{a} \mapsto \boldsymbol{b}$ 是线性的. 它的矩阵是什么?

3.11 给定 $\mathrm{Hom}(\mathbb{R}^2)$ 中的 S 与 T, 它们的矩阵分别是

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad 与 \quad \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix},$$

试找出 $\mathrm{Hom}(\mathbb{R}^2)$ 中 $S \circ T$ 的矩阵.

3.12 给定 $\mathrm{Hom}(\mathbb{R}^2)$ 中的 S 与 T, 它们的矩阵为

$$s = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} \quad 与 \quad t = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix},$$

试找出 $S \circ T$ 的矩阵.

3.13 将上面的解答记在心上, 如果 S 与 T 在 $\mathrm{Hom}(\mathbb{R}^3)$ 之中时, 你会猜想 $S \circ T$ 的矩阵是什么? 试验证你的猜想.

3.14 我们知道若 $T \in \mathrm{Hom}(V, W)$ 为一个同构, 则 T^{-1} 为 $\mathrm{Hom}(W, V)$ 中一个同构, 试证

$$S \circ T满射 \Rightarrow S满射, \quad S \circ T单射 \Rightarrow T单射,$$

从而如果 $T \in \mathrm{Hom}(V, W), S \in \mathrm{Hom}(W, V)$, 及

$$S \circ T = I_V, \quad T \circ S = I_W,$$

则 T 是一个同构.

3.15 试证若 S^{-1} 与 T^{-1} 存在, 则 $(S \circ T)^{-1}$ 存在并等于 $T^{-1} \circ S^{-1}$, 试给出这个结果一个更细致的语句.

3.16 试证如果 $\mathrm{Hom}\, V$ 中的 S 与 T 是可交换的, 则 T 的零空间 $N = N(T)$ 与它的值域 $R = R(T)$ 在 S 之下是不变的 $(S[N] \subset N$ 及 $S[R] \subset R)$.

3.17 试证如果 α 是 T 的一个特征向量及 S 与 T 是可交换的, 则 $S(\alpha)$ 是 T 的一个特征向量并有同样的特征值.

3.18 试证如果 S 与 T 是可交换的及 T^{-1} 存在, 则 S 与 T^{-1} 可以交换.

3.19 给定 T 的一个特征向量 α, 它有特征值 x, 试证 α 亦为 $T^2 = T \circ T, T^n$, 及 T^{-1}(如果 T 是可逆的) 的一个特征向量, 而且对应的特征值为 x^2, x^n, 及 $1/x$.

给定 t 的一个多项式 $p(t)$, 定义算子 $p(T)$, 及在上面假设之下, 试证 α 是 $p(T)$ 的一个特征向量且有特征值 $p(x)$.

3.20 若 S 与 T 在 $\mathrm{Hom}\, V$ 之中, 如果 $\mathrm{Hom}\, V$ 中每一个可与 T 交换的 A 都可与 S 交换, 则我们称 S 与 T 是二重交换的 (记为 $S \mathrm{cc} T$). 固定 T, 置 $\{T\}'' = \{S : S\mathrm{cc}T\}$. 试证 $\{T\}''$ 是 $\mathrm{Hom}\, V$ 的一个交换子代数.

3.21 给定 Hom$V$ 中的 T 及 $\alpha\in V$, 命 N 为 "T 之下, α 的轨道" (集合 $\{T^n\alpha:n\in\mathbb{Z}^+\}$) 的线性生成. 试证在 T 之下, N 是不变的.

3.22 一个 Hom$V$ 中的变换 T 满足 $T^n=0$ 对于某 n 成立, 则称为幂零的.试证若 T 是幂零的, 则 $I-T$ 是可逆的.[提示: 若将 x 换成 T, 则幂级数

$$\frac{1}{1-x}=\sum_{0}^{\infty}x^n$$

是一个有限和.]

3.23 假定 T 是幂零的, S 与 T 可交换, 及 S^{-1} 存在, 其中 $S,\ T\in$ Hom$V$. 试证 $(S-T)^{-1}$ 存在.

3.24 命 φ 为由一个向量空间 V 至一个向量空间 W 的同构. 试证 $T\to\varphi\circ T\circ\varphi^{-1}$ 是一个由代数 Hom$V$ 至代数 Hom$W$ 的代数同构.

3.25 试利用停止箭头记号证明关于 $\mathbb{R}^3=\mathbb{R}\times\mathbb{R}\times\mathbb{R}$ 的所有 π_i 与 θ_i 是显式的. 同时将恒等式 $\sum\theta_j\circ\pi_j=I$ 写成显形式.

3.26 对于 $\mathbb{R}^5=\mathbb{R}^2\times\mathbb{R}^3$ 做同样的事.

3.27 试证前两个投影单射恒等式 ($\pi_i\circ\theta_i=I_i$ 及当 $j\neq i$ 时, $\pi_j\circ\theta_i=0$) 只是 θ_i 定义的一个重述. 再证从这些恒等式与定理 3.4 形式地推出 θ_i 的线性性.

3.28 试证恒等式 $\sum\theta_i\circ\pi_i=I$. 将 π_j 作用于这个方程并回忆一下, 若对于所有 j 皆有 $\pi_j(f)=\pi_j(g)$ (这正好就是对于所有 j, 方程 $f(j)=g(j)$), 则 $f=g$.

3.29 试证定理 3.6 的一般情形. 我们给出有共同定义域 V 及值域 $\{W_i:i\in I\}$ 的线性映射的一个有指标集合 $\{T_i:i\in I\}$. 第一个问题为如何去定义 $T:V\to W=\prod_i W_i$. 对于每个 $\xi\in V$, 由适当地定义 $T(\xi)$ 来做这件事, 然后利用定理 3.4 来得出 T 是线性的.

3.30 试证定理 3.7.

3.31 未作计算, 我们就知道由 $\mathbb{R}^3$ 至 $\mathbb{R}^4$ 的映射

$$T:\boldsymbol{x}\to\prec 3x_1-x_2+x_3,x_2+x_3,x_1-5x_3,2x_1\succ$$

是线性的. 为什么?(从课文中引用相关的定理.)

3.32 试写出上面例子中的变换 T 之矩阵, 然后以显式有序四元形式写出由 $\mathbb{R}$ 至 $\mathbb{R}^4$ 的映射 $T\circ\theta_i(i=1,2,3)$.

3.33 命 $W=\prod_1^n W_i$ 为向量空间的有限乘积并置 $p_i=\theta_i\circ\pi_i$, 所以对于所有 i, p_i 皆在 Hom$W$ 之中. 由投影 - 单射恒等式证明 $\sum_1^n p_i=I$(W 上的恒等映射), 当 $i\neq j$ 时, $p_i\circ p_j=0$, 及 $p_i\circ p_i=p_i$. 值域 $R_i=R(p_i)$ 视为等同.

3.34 在上面习题中, 定义 Hom W 中的 T 为

$$\sum_{m=1}^{n} mp_m.$$

试证 α 为 T 的一个特征向量当且仅当 α 在一个子空间 R_i 中, 及 α 的特征值是 i.

3.35 在同样的情况下, 试证多项式

$$\prod_{j=1}^{n}(T-jI)=(T-I)\circ\cdots\circ(T-nI)$$

为零变换.

3.36 若 $T\in \mathrm{Hom}(V, W)$, 则定理 3.6 与 3.7 可以被联合起来, 此处 V 与 W双方都是乘积空间:

$$V=\prod_{1}^{n} V_j \quad 与 \quad W=\prod_{1}^{m} W_i.$$

试陈述并证明一条定理: 一个 T 可以被分解为一个二重有指标的系 $\{T_{ij}\}$, 其中 $T_{ij}\in \mathrm{Hom}(V_i, W_j)$ 及反之, 任意这种二重有指标的系皆可以被聚合起来形成一个由 V 至 W 的单个 T.

3.37 将你的定理用于特殊情况, 此处 $V=\mathbb{R}^n$ 及 $W=\mathbb{R}^m$(即对于所有 i 及 j, $V_i=W_j=\mathbb{R}$). 现在 T_{ij} 由 $\mathbb{R}$ 至 $\mathbb{R}$, 所以它仅仅是一个数 t_{ij} 的乘积. 试证这些数的有指标集合 $\{t_{ij}\}$ 为 T 的矩阵.

3.38 给定任意向量空间的 m 元组 $\{W_i\}_1^m$, 假定有一个向量空间 X 及 $\mathrm{Hom}(X, W_i)$ 中的映射 $p_i, i=1,\cdots,m$, 具有下面性质:

P. 对于任意从一个共同定义空间 V 至上述空间 W_i 的线性映射 m 元组 $\{T_i\}$ (所以 $T_i\in \mathrm{Hom}(V, W_i), i=1,\cdots,m$), 皆存在 $\mathrm{Hom}(V, X)$ 中惟一的 T 使 $T_i=P_i\circ T, i=1,\cdots,m$.

试证存在一个由

$$W=\prod_{1}^{m} W_i \quad 至 \quad X$$

的 "标准的" 同构, 在它之下, 给予的映射 p_i 变成了投影 π_i[注记: 由定理 3.6 可知乘积空间 W 本身具有性质 P, 所以这个习题表示 P 是乘积空间的一个抽象的特征化].

1.4 仿射子空间与商空间

在这一节中, 我们将关注一个向量空间 V 中的 "平面" 及当我们平移它们, 将它们互相相交, 取其线性映射之像, 等等之后, 看看对它们发生了什么. 然后我

们将专注于一个固定子空间的所有平移, 即所有平面之集合并发现这个集合自身以最平凡的方式为一个向量空间. 这个材料中的一部分已放在第 2 节之中.

仿射子空间 若 N 是一个向量空间 V 的一个子空间, α 为 V 的任意向量, 则集合 $N+\alpha=\{\xi+\alpha:\xi\in N\}$ 或者被称为 N 包含 α 的陪集,或者为 V 的过 α 且平行于 N 的仿射子空间. 集合 $N+\alpha$ 亦被称为 N 通过 α 的平移.我们在第 2 节中看到仿射子空间是我们想称之为平面的一般物体. 若在一个讨论中, 给出固定的 N, 我们将使用记号 $\overline{\alpha}=N+\alpha$(见 0.12 节).

我们开始于列举仿射子空间的某些简单性质, 其中一些将推广第 2 节中已作过的考察, 并将一些证明留作习题.

(1) 假定 N 是一个固定的子空间, 若 $\gamma\in\overline{\alpha}$, 则 $\overline{\gamma}=\overline{\alpha}$. 事实上, 若 $\gamma=\alpha+\eta_0$, 则 $\gamma+\eta=\alpha+(\eta_0+\eta)\in\overline{\alpha}$, 所以 $\overline{\gamma}\subset\overline{\alpha}$. 同时, $\alpha+\eta=\gamma+(\eta-\eta_0)\in\overline{\gamma}$, 所以 $\overline{\alpha}\subset\overline{\gamma}$. 因此 $\overline{\alpha}=\overline{\gamma}$.

(2) 对于固定 N, 给予任意 α 与 β, 则或者 $\overline{\alpha}=\overline{\beta}$, 或者 $\overline{\alpha}$ 与 $\overline{\beta}$ 不相交. 事实上, 若 $\overline{\alpha}$ 与 $\overline{\beta}$ 不是不相交, 则存在 γ 属于每一个, 从而由 (1) 可知 $\overline{\alpha}=\overline{\gamma}=\overline{\beta}$. 读者可以发现它启发我们去比较这些计算与 0.12 节中较一般的计算. 在此 $\alpha\sim\beta$ 当且仅当 $\alpha-\beta\in N$.

(3) 现在命 $\mathfrak{a}$ 为 V 的所有仿射子空间的集合; 因此 $\mathfrak{a}$ 为 V 的所有向量子空间的所有陪集之集合。则 $\mathfrak{a}$ 的任意子系之交或为空集或自身为一个仿射子空间. 事实上, 若 $\{A_i\}_{i\in I}$ 为仿射子空间的一个有指标集合, A_i 为向量子空间 W_i, 对于每一个 $i\in I$ 的一个陪集, 则 $\bigcap_{i\in I}A_i$ 或者为空的或者为向量子空间 $\bigcap_{i\in I}W_i$ 的一个陪集. 事实上, 若 $\beta\in\bigcap_{i\in I}A_i$, 则由 (1) 可知对于所有 i, $A_i=\beta+W_i$, 从而 $\bigcap A_i=\beta+\bigcap W_i$.

(4) 若 $A,B\in\mathfrak{a}$, 则 $A+B\in\mathfrak{a}$. 即任意两个仿射子空间的集和自身就是一个仿射子空间.

(5) 若 $A\in\mathfrak{a}$ 及 $T\in\mathrm{Hom}(V,W)$, 则 $T[A]$ 是 W 的一个仿射子空间. 特别, 若 $t\in\mathbb{R}$, 则 $tA\in\mathfrak{a}$.

(6) 若 B 是 W 的一个仿射子空间及 $T\in\mathrm{Hom}(V,W)$, 则 $T^{-1}[B]$ 或为空的或为 V 的一个仿射子空间.

(7) 对于一个固定的 $\alpha\in V$, V经过 α 的平移为由 $S_\alpha(\xi)=\xi+\alpha$, 对所有 $\xi\in V$, 定义的映射 $S_\alpha:V\to V$. 平移不是线性的. 例如, $S_\alpha(0)=\alpha$. 但是, 这是很清楚的, 即平移将仿射子空间映入仿射子空间. 因此, $S_\alpha(A)=A+\alpha$ 及 $S_\alpha(\beta+W)=(\alpha+\beta)+W$.

(8) 由一个向量空间 V 至一个向量空间 W 的仿射变换是一个 V 至 W 的

线性映射再接着一个 W 中的平移. 因此一个仿射变换具有形式 $\xi \mapsto T(\xi)+\beta$, 此处 $T \in \mathrm{Hom}(V, W)$ 及 $\beta \in W$. 注意 $\xi \mapsto T(\xi+\alpha)$ 是仿射变换, 这是由于

$$T(\xi+\alpha)=T(\xi)+\beta, \quad \text{此处} \quad \beta=T(\alpha).$$

由 (5) 与 (7) 可知一个仿射变换将 V 的仿射子空间映至 W 的仿射子空间.

商空间 现在固定 V 的一个子空间 N, 并考虑所有 N 的平移 (陪集) 的集合 W. 我们往证, 在可能最自然的方式下, W 自身就是一个向量空间. 加法将是集合加法, 及标量乘法将是集合乘法 (除去一个特殊情况). 例如, 若 N 为 $\mathbb{R}^3$ 中过原点的一条直线, 则 W 包含 $\mathbb{R}^3$ 中所有平行于 N 的直线. 我们说这个平行直线所成之集合将自动地是一个向量空间: W 中任何两条直线之集合和仍为 W 中一条直线! 又若 $L \in W$ 及 $t \neq 0$, 则集合乘积 tL 是 W 中的一条直线. L 的平移纤维了 $\mathbb{R}^3$, 而纤维之集合是一个自然的向量空间.

在这个讨论中, 我们暂时地用 '$+_s$' 表示集合和及 '$\cdot_s$' 表示集合乘积是有助益的. 当 N 固定时, 由上面的 (2) 可知两个陪集或者不相交或者恒等. 所以, 在一般情况下, 所有陪集所成之集合 W 是 V 的一个纤维表示, 恰如我们关于平行直线的例子一样. 由 (4) 或由一个直接的计算可知 $\overline{\alpha}+_s\overline{\beta}=\overline{\alpha+\beta}$. 因此 W 在加法之下是闭的, 自然地, 我们取这个作为 W 上的加法, 即我们由 $\overline{\alpha}+\overline{\beta}=\overline{\alpha}+_s\overline{\beta}$ 来定义 W 上的 $+$. 则由 V 至 W 的自然映射 $\pi: \alpha \rightarrow \overline{\alpha}$ 保持加法; $\pi(\alpha+\beta)=\pi(\alpha)+\pi(\beta)$, 这是由于这正好是我们的上面的方程 $\overline{\alpha+\beta}=\overline{\alpha}+\overline{\beta}$. 类似地, 若 $t \in \mathbb{R}$, 则集合乘积 $t \cdot_s \overline{\alpha}$ 或为 $\overline{t\alpha}$ 或为 $\{0\}$. 所以, 如果当 $t \neq 0$ 及当 $t=0$ 时 $\overline{0}=N$, 我们定义 $t\overline{\alpha}$ 为集合乘积, 则 π 亦保持标量乘法,

$$\pi(t\alpha)=t\pi(\alpha).$$

因此我们有 N 的所有陪集之集合 W 上有两个向量类似运算, 从而我们自然地盼望 W 是一个向量空间. 我们可以由验证所有定律来证明这个事实, 但是去关注这样一个验证证明的一般方式将更加优美.

定理 4.1 命 V 为一个向量空间及 W 为一个具有两个我们通常称呼的向量类似运算之集合. 假定存在一个满射映射 $T: V \rightarrow W$ 保持这些运算:$T(s\alpha+t\beta)=sT(\alpha)+tT(\beta)$, 则 W 是一个向量空间.

证明 我们来验证定律 A1~S4. 无论如何, 举一个例子应该使读者清楚如何来进行. 我们欲证 $T(0)$ 满足 A3, 从而它是 W 的零向量. 由于每一个 $\beta \in W$ 都

具有形式 $T(\alpha)$, 所以

$$T(0)+\beta=T(0)+T(\alpha)=T(0+\alpha)=T(\alpha)=\beta.$$

这就是 A3. 我们将要读者在习题中验证更多的定律. □

定理 4.2 一个向量空间 V 的一个固定 N 的陪集之集合本身就构成一个向量空间, 称为上述自然运算下的商空间 V/N, 而且投影 π 是一个由 V 至 V/N 的满射线性映射.

定理 4.3 若 T 属于 $\operatorname{Hom}(V,W)$ 及若 T 的零空间包含子空间 $M\subset V$, 则经过 V/M, T 有一个惟一的分解. 即在 $\operatorname{Hom}(V/M,W)$ 中存在一个惟一的变换 S 使 $T=S\circ\pi$.

证明 由于 T 在 M 上为零, 所以 T 在 M 的每一个陪集 A 上为常数, 所以 $T[A]$ 仅含有一个向量. 如果我们定义 $S(A)$ 为 $T[A]$ 中惟一的向量, 则 $S(\overline{\alpha})=T(\alpha)$, 所以由定义得 $S\circ\pi=T$. 反之, 若 $T=R\circ\pi$, 则 $R(\overline{\alpha})=R\circ\pi(\alpha)=T(\alpha)$, 从而 R 为我们上述之 S. S 的线性性实际上是显然的. 因此

$$S(\overline{\alpha}+\overline{\beta})=S(\overline{\alpha+\beta})=T(\alpha+\beta)=T(\alpha)+T(\beta)=S(\overline{\alpha})+S(\overline{\beta}),$$

类似地, 可得齐性. 定理证完. □

再给一个在此有兴趣的注记. 若 N 在 $\operatorname{Hom}V$ 中的一个线性映射之下是不变的 (即 $T[N]\subset N$), 则对于每一个 V 中之 α, $T[\overline{\alpha}]$ 是陪集 $\overline{T(\alpha)}$ 的一个子集合, 这是由于

$$T[\overline{\alpha}]=T[\alpha+N]=T(\alpha)+_s T[N]\subset T(\alpha)+_s N=\overline{T(\alpha)}.$$

所以有一个映射 $S:V/N\to V/N$, 它由要求 $S(\overline{\alpha})=\overline{T(\alpha)}$ 来定义 (或 $S\circ\pi=\pi\circ T$), 而且易于验证 S 是线性的. 故得

定理 4.4 若一个向量空间 V 的一个向量子空间 N 被 $\operatorname{Hom}V$ 中的一个变换 T 映入自身, 则在 $\operatorname{Hom}(V/N)$ 中存在一个惟一的变换 S 使 $S\circ\pi=\pi\circ T$.

习题

4.1 试证仿射子空间的性质 (4),(5) 与 (6).

4.2 在欧氏平面 $\mathbb{E}^2$ (即你手中的纸) 中选取一个原点 O, 并命 L_1 与 L_2 为两条不包含 O 的直线. 命 X 与 Y 为 L_1 上相异的点及 Z 为 L_2 上的任意点. 试画出表示几何和

$$\overrightarrow{OX}+\overrightarrow{OZ} \quad 与 \quad \overrightarrow{OY}+\overrightarrow{OZ}$$

的图 (平行四边形法则), 并表述来自平面几何的定理, 即这两个和点位于平行于 L_1 与 L_2 的第三条直线 L_3 上.

4.3 (a) 试证关于定理 4.1 的加法结合律.
(b) 试证定律 A4 与 S2.

4.4 现在回到习题 2.1, 并按照定理 4.1 重新检查一下, 最后试证我们如何真正知道几何向量构成一个向量空间.

4.5 试证定理 4.3 中的映射 S 是单射当且仅当 N 是 T 的零空间.

4.6 由习题 4.5, 我们知道如果 T 是 $\mathrm{Hom}(V, W)$ 中的一个满射元素及 N 为 T 的零空间, 则定理 4.3 中的 S 是一个由 V/N 至 W 的同构. 其逆 S^{-1} 对 W 的每一个 η 皆使之对应 N 的一个陪集. 试证 "不定积分" 的程序是这样一种映射 S^{-1} 的一个例子. 这是计算一个积分并加上一个任意常数的程序, 如同

$$\int \sin x dx = -\cos x + c.$$

4.7 假定 N 与 M 为一个向量空间 V 的子空间及 $N \subset M$. 试证 M/N 是 V/N 的一个子空间及 V/M 自然地同构于商空间 $(V/N)/(M/N)$. [提示: N 的每一个陪集是 M 的某陪集之子集合.]

4.8 假定 N 与 M 为一个向量空间 V 的任意子空间. 试证 $(M+N)/N$ 自然地同构于 $M/(M\bigcap N)$.(由这样的事实出发, 即 $M\bigcap N$ 的每一个陪集皆含于 N 的一个惟一的陪集之中.)

4.9 试证定理 4.4 中的映射 S 是线性的.

4.10 给予 $T\in \mathrm{Hom}\,V$, 试证 $T^2=0(T^2=T\circ T)$ 当且仅当 $R(T)\subset N(T)$.

4.11 假定 $T\in \mathrm{Hom}\,V$ 及子空间 N 满足: T 在 N 上与 V/N 上为恒等映射. 后面的假定为定理 4.4 中的 S 在 V/N 上为恒等映射. 置 $R=T-I$, 并用上面习题证明 $R^2=0$. 再证明若 $T=I+R$ 及 $R^2=0$, 则存在一个子空间 N 使 T 在 N 上及在 V/N 上都是恒等映射.

4.12 现在我们略为不同地观察上面的情况. 假定 T 在 N 与 V/N 上是恒等的, 置 $R=I-T$, 试证存在一个 $K\in \mathrm{Hom}(V/N, V)$ 使 $R=K\circ\pi$. 再证对于任意 N 的陪集 A, T 在 A 上的作用可以被看作经过 $K(A)$ 的平移. 即若 $\xi\in A$ 及 $\eta=K(A)$, 则 $T(\xi)=\xi+\eta$.

4.13 考虑 HomR^2 中的映射 $T: \prec x_1, x_2 \succ \mapsto \prec x_1 + 2x_2, x_2 \succ$ 及命 N 为 $R = T - I$ 的零空间. 等同 N 并证 T 在 N 与 $\mathbb{R}^2/N$ 上是恒等的. 试找出上面习题中的映射 K. 这样一个映射 T 被称为 V 平行于 N 的*剪状变换*. 试画出单位平方及它在 T 之下的像.

4.14 如果我们回忆一下, 一个向量空间 V 的一个子集合 A 的线性生成 $L(A)$ 可以被定义为 V 的所有包含 A 的子空间的交, 则一个向量空间 V 的仿射子空间的任意集合之交或者为一个仿射子空间或者为空集. 这个事实建议我们定义一个非空子集合 $A \subset V$ 的*仿射生成*$M(A)$ 为所有包含 A 的仿射子空间之交. 则由我们列举的仿射性质 (3) 得知 $M(A)$ 是一个仿射子空间, 而且由上面的定义可知, 它是包含 A 的最小仿射子空间. 现在我们自然地惊奇, $M(A)$ 是不是能够直接地用线性组合的词句来加以描述. 试首先证明若 $\alpha \in A$, 则 $M(A) = L(A - \alpha) + \alpha$; 然后证 $M(A)$ 是 A 上满足 $\sum x_i = 1$ 的所有线性组合 $\sum x_i\alpha_i$ 所成之集合.

4.15 试证一个集合 B 的线性生成是 $B \cup \{0\}$ 的仿射生成.

4.16 试证对于 V 中的任意 γ 有 $M(A + \gamma) = M(A) + \gamma$ 及对于任何 $x \in \mathbb{R}$ 有 $M(xA) = xM(A)$.

1.5 直和

我们现在转入本章之中心. 经常会发生这种情况, 即一个向量空间 V 的某些现象之研究导致一个子空间的有限集合 $\{V_i\}$ 使 V 自然地同构于乘积空间 $\prod_i V_i$. 在这个同构之下, 乘积空间上的映射 $\theta_i \circ \pi_i$ 变成 Hom$V$ 中某些映射 P_i, 及投影 - 单射恒等式被反映为恒等式 $\sum P_i = I$, 对于所有 j, $P_j \circ P_j = P_j$, 及当 $i \neq j$ 时, $P_i \circ P_j = 0$. 同时, $V_i = P_i$ 的值域. V 这样得到的乘积结构即被用来研究对它产生的现象. 例如这是我们揭示 Hom$V$ 中一个线性变换的结构之途径, 对它的研究是线性代数的中心问题之一.

直和 若 $V_1, \cdots, V_n$ 是向量空间 V 的子空间, 则映射 $\pi: \prec \alpha_1, \cdots, \alpha_n \succ \mapsto \sum_1^n \alpha_i$ 是从 $\prod_1^n V_i$ 至 V 的一个线性变换, 这是由于它是坐标投影之和 $\pi = \sum_1^n \pi_i$.

定义 5.1 *若 π 为单射, 我们将称 V_i 为独立的, 及若 π 是一个同构, 则 V 为所有 V_i 之直和. 我们将后面的关系表示为 $V = V_1 \bigoplus \cdots \bigoplus V_n = \bigoplus_1^n V_i$.*

因此 $V = \bigoplus_{i=1}^n V_i$ 当且仅当 π 是单射与满射的, 即当且仅当子空间 $\{V_i\}_1^n$ 为独立的, 并生成 V. 直和条件的一个有用之重述为: 每一个 $\alpha \in V$ 惟一地表示为一个和 $\sum_1^n \alpha_i$, 此处对于所有 i, $\alpha_i \in V_i$; 因为这些 V_i 生成 V, 所以 α 有一些这种表示, 及由于它们的独立性, 所以表示是惟一的.

例如, 命 $V=\mathfrak{C}(\mathbb{R})$ 为 $\mathbb{R}$ 上的实值连续函数空间, 命 V_e 为偶函数的子集合 (对于所有 x, 适合 $f(-x)=f(x)$ 之 f), 及命 V_o 为奇函数之子集合 (对于所有 x, 适合 $f(-x)=-f(x)$ 之 f). 显然 V_e 与 V_o 都是 V 的子空间. 我们宣布 $V=V_e\bigoplus V_o$. 欲证明这一点, 注意对于 V 中任意 f, $g(x)=(f(x)+f(-x))/2$ 为偶的, $h(x)=(f(x)-f(-x))/2$ 为奇的, 及 $f=g+h$. 因此 $V=V_e+V_o$. 进而言之, f 的这个分解是惟一的. 事实上, 若 $f=g_1+h_1$, 此处 g_1 为偶的及 h_1 为奇的, 则 $g-g_1=h_1-h$. 由于既为偶又为奇的仅有之函数为零, 所以 $g-g_1=0=h_1-h$. e^x 的偶 - 奇分量为双曲正弦与双曲余弦函数:

$$e^x=\frac{(e^x+e^{-x})}{2}+\frac{(e^x-e^{-x})}{2}=\cosh x+\sinh x.$$

由于 π 为单射当且仅当它的零空间为 $\{0\}$(引理 1.1), 所以我们有:

引理 5.1 子空间 $\{V_i\}_1^n$ 的独立性等价于这样的性质, 即若对于所有 $i,\alpha_i\in V_i$ 与 $\sum_1^n\alpha_i=0$, 则对于所有 $i,\alpha_i=0$.

推论 5.1 若子空间 $\{V_i\}_1^n$ 是独立的, 对于所有 i, $\alpha_i\in V_i$, 及 $\sum_1^n\alpha_i$ 为 V_j 中的一个元素, 则当 $i\neq j$ 时, $\alpha_i=0$.

我们将证明留给读者去做.

两个子空间的情况特别简单

引理 5.2 V 的子空间 M 与 N 是独立的当且仅当 $M\cap N=\{0\}$.

证明 若 $\alpha\in M,\beta\in N$, 及 $\alpha+\beta=0$, 则 $\alpha=-\beta\in M\cap N$. 如果 $M\cap N=\{0\}$, 则这将进一步推出 $\alpha=\beta=0$, 所以 M 与 N 是独立的. 另一方面, 若 $0\neq\beta\in M\cap N$, 又若我们置 $\alpha=-\beta$, 则 $\alpha\in M,\beta\in N$, 及 $\alpha+\beta=0$, 所以 M 与 N 是不独立的. □

注意上面的第一个论证仅为我们早先给出的, $\mathbb{R}$ 上一个函数的偶 - 奇分解的惟一性论证的一般形式.

推论 5.2 $V=M\bigoplus N$ 当且仅当 $V=M+N$ 及 $M\cap N=\{0\}$.

定义 5.2 若 $V=M\bigoplus N$, 则 M 与 N 称为补子空间, 并称每一个为另一个之补.

警示 V 的一个子空间 M, 除非是平凡的 (即 $M=\{0\}$ 或 $M=V$), 它没有一个惟一的补子空间. 如果我们将 $\mathbb{R}^3$ 看成坐标化的欧氏三维空间, 则 M 是一个真子空间当且仅当 M 为包含原点的一张平面或 M 为过原点的一条直线 (见图 1.9). 如果 M 与 N 为真子空间. 其中一个为一张平面及另一个为一条不位于该平面上的直线, 则 M 与 N 为补子空间. 此外, 这些是 $\mathbb{R}^3$ 中仅有的非平凡补子空间对. 读者将被要求在习题中证明某些事实, 而在下一章之中间部分, 它们都将会清楚的.

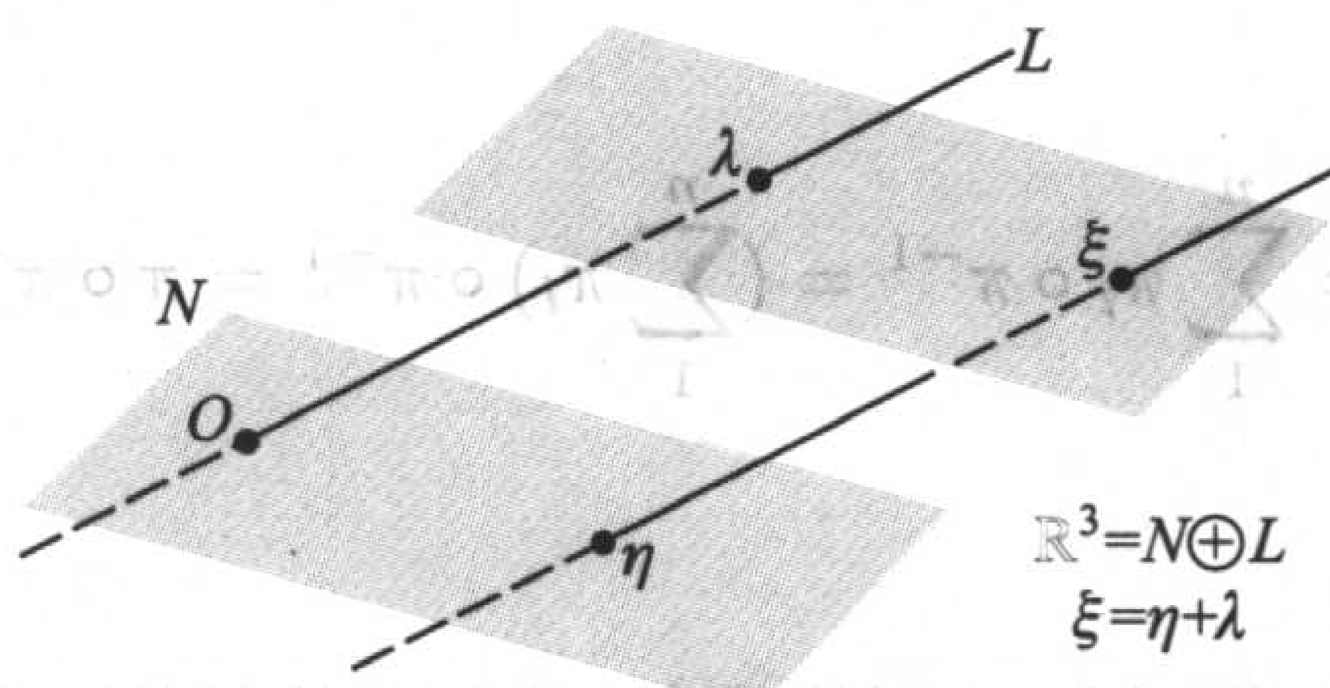


图 1.9

下面的引理在技术上是有用的.

引理 5.3 若 V_1 与 V_0 为 V 的独立子空间及 $\{V_i\}_2^n$ 为 V_0 的独立子空间, 则 $\{V_i\}_1^n$ 为 V 的独立子空间.

证明 若对于所有 $i, \alpha_i \in V_i$, 及 $\sum_1^n \alpha_i = 0$, 则置 $\alpha_0 = \sum_2^n \alpha_i$, 我们有 $\alpha_1 + \alpha_0 = 0$, 其中 $\alpha_0 \in V_0$. 所以由 V_1 与 V_0 的独立性可知 $\alpha_1 = \alpha_0 = 0$. 然后由 $\{V_i\}_2^n$ 的独立性可知 $\alpha_2 = \alpha_3 = \cdots = \alpha_n = 0$, 引理证完 (引理 5.1). □

推论 5.3 $V = V_1 \bigoplus V_0$ 与 $V_0 = \bigoplus_{i=2}^n V_i$ 一起可以导出 $V = \bigoplus_{i=1}^n V_i$.

投影 如果 $V = \bigoplus_{i=1}^n V_i$, 若 π 为同构 $\prec \alpha_1, \cdots, \alpha_n \succ \mapsto \alpha = \sum_1^n \alpha_i$, 及若 π_j 为由 $\prod_{i=1}^n V_i$ 至 V_j 的第 j 个投影映射 $\prec \alpha_1, \cdots, \alpha_n \succ \mapsto \alpha_j$, 则 $(\pi_j \circ \pi^{-1})(\alpha) = \alpha_j$.

定义 5.3 我们称 α_j 为 α 的第 j 个分量,及我们称线性映射 $P_j = \pi_j \circ \pi^{-1}$ 为 V 至 V_j 上的投影(关于 V 的给定直和分解). 由于 V 中每一个 α 惟一地可表为一个和 $\alpha = \sum_1^n \alpha_i$, 其中对于所有 i, $\alpha_i \in V_i$, 所以我们可以将 $P_j(\alpha) = \alpha_j$ 看作"α 在 V_j 中的部分".

"投影" 这个字在这里的使用与它在笛卡儿乘积情况的使用是不同的, 而每一个又与它在商空间范围中的使用相异 (0.12 节). 显然这三种应用是相关联的, 因为真正的含义在上下文中总是清楚的, 从而含糊性只会引起小的混乱.

定理 5.1 若映射 P_i 为上面的投影, 则值域 $P_i = V_i$, 当 $i \neq j$ 时, $P_i \circ P_j = 0$, 及 $\sum_1^n P_i = I$.

证明 由于 π 为一个同构及 $P_j = \pi_j \circ \pi^{-1}$, 所以值域 $P_j=$ 值域 $\pi_j = V_j$. 其次, 由引理 5.1 的推论直接推出, 若 $\alpha \in V_j$, 则 $P_i(\alpha) = 0(i \neq j)$, 从而 $P_i \circ P_j = 0(i \neq j)$. 最后,

$$\sum_1^n P_i = \sum_1^n \pi_i \circ \pi^{-1} = (\sum_1^n \pi_i) \circ \pi^{-1} = \pi \circ \pi^{-1} = I,$$

定理证完. □

上面的投影性质显然是关于同构空间 $\prod_1^n V_i$ 的投影 - 单射恒等式在 V 中的反映.

一个逆定理亦成立.

定理 5.2 若 $\{P_i\}_1^n \subset \operatorname{Hom} V$ 满足 $\sum_1^n P_i = I$ 与 $P_i \circ P_j = 0(i \neq j)$, 及若我们置 $V_i=$ 值域 P_i, 则 $V = \bigoplus_{i=1}^n V_i$, 及 P_i 为 V_i 上的对应投影.

证明 方程 $\alpha = I(\alpha) = \sum_1^n P_i(\alpha)$ 表明了子空间 $\{V_i\}_1^n$ 生成 V. 其次, 若 $\beta \in V_j$, 则由于 $\beta \in$ 值域 P_j 及 $P_i \circ P_j = 0(i \neq j)$, 所以 $P_i(\beta) = 0(i \neq j)$. 我们亦有 $P_j(\beta) = (I - \sum_{i \neq j} P_i)(\beta) = I(\beta) = \beta$. 现在对于任意选取的 $\alpha_i \in V_i$, 我们考虑 $\alpha = \sum_1^n \alpha_i$. 利用上面两个事实, 我们有 $P_j(\alpha) = P_j(\sum_{i=1}^n \alpha_i) = \sum_{i=1}^n P_j(\alpha_i) = \alpha_j$. 所以, 由 $\alpha = 0$ 推出 $\alpha_j = P_j(0) = 0$ 对于所有 j 成立, 及诸子空间 V_i 是独立的. 从而 $V = \bigoplus_1^n V_i$. 最后, $\alpha = \sum P_i(\alpha)$ 及对于所有 i, $P_i(\alpha) \in V_i$ 这个事实表示对于每一个 $\alpha, P_j(\alpha)$ 是 α 的第 j 个分量, 所以 P_j 是 V 在 V_j 上的投影. □

存在一个映射类的内蕴特征化, 即一个投影.

引理 5.4 投影 P_i 为幂等的 ($P_i^2 = P_i$), 或等价地, 每一个在其值域上都是恒等的. P_i 的零空间为空间 $V_j(j \neq i)$ 之和.

证明 $P_j^2 = P_j \circ (I - \sum_{i \neq j} P_i) = P_j \circ I = P_j$. 由于这可以被改写为, 对于每一个 V 中的 α, $P_j(P_j(\alpha)) = P_j(\alpha)$, 这恰好是说 P_j 是其值域的恒等映射.

现在置 $W_i = \sum_{j \neq i} V_j$, 并注意到若 $\beta \in W_i$, 则 $P_i(\beta) = 0$, 这是由于 $P_i[V_j] = 0 (j \neq i)$. 因此 $W_i \subset N(P_i)$. 反之, 若 $P_i(\alpha) = 0$, 则 $\alpha = I(\alpha) = \sum_1^n P_j(\alpha) = \sum_{j \neq i} P_j(\alpha) \in W_i$. 因此 $N(P_i) \subset W_i$, 从而这两个空间相等. □

反之:

引理 5.5 若 $P \in \mathrm{Hom}(V)$ 是幂等的, 则 V 是它的值域与零空间的直和, 及 P 为其值域上的对应投影.

证明 置 $Q = I - P$, 我们有 $PQ = P - P^2 = 0$. 所以由上面定理可知 V 是 P 与 Q 的值域的直和, 及 P 为其值域上的对应投影. 进而言之, 由推论可知 Q 的值域为 P 的零空间. □

若 $V = M \bigoplus N$ 及 P 为在 M 上的对应投影, 我们称 P 为M 上沿 N 的投影. 由于 M 不决定 N, 所以投影 P 不被 M 单独决定. $\mathrm{Hom} V$ 中满足 $P + Q = I$ 及 $PQ = QP = 0$ 的一对 P 与 Q 被称为一对补投影.

在上面的讨论中, 我们略去了另一细微点. 严格地讲, 当我们形成和 $\pi = \sum_1^n \pi_i$ 时, 我们处理每一个 π_i 好像它是由 $\prod_1^n V_i$ 至 V, 其实真正地, π_j 的上域为 V_j. 我们要求 P_j 是由 V 到 V, 其实 $\pi_j \circ \pi^{-1}$ 有值域 V_j, 所以方程 $P_j = \pi_j \circ \pi^{-1}$ 不能很精确. 要修复这些缺陷, 我们引进单射映射 $\iota_j : V_j \to V$, 它在 V_j 上是恒等映射, 但它将 V_j 看成 V 的一个子空间, 所以取 V 为其上域. 若我们关于一个映射的概念包含一个比上域尽可能大的一个值域, 则我们可以允许这种恒等单射. 所以, 置 $\bar{\pi}_j = \iota_j \circ \pi_j$, 则我们有正确的方程 $\pi = \sum_1^n \bar{\pi}_i$ 与 $P_j = \bar{\pi}_j \circ \pi^{-1}$.

习题

5.1 试证引理 5.1 的推论

5.2 命 α 为 $\mathbb{R}^3$ 中之向量 $\prec 1, 1, 1 \succ$, 及命 $M = \mathbb{R}\alpha$ 为它的一维生成. 试证三个坐标平面中的每一个都是 M 的补.

5.3 试证一个有限积空间 $V = \prod_1^n V_i$ 有子空间 $\{W_i\}_1^n$ 使 W_i 同构于 V_i 和 $V = \bigoplus_1^n W_i$. 并指出对应的投影 $\{P_i\}$ 如何相关于 π_i 与 θ_i.

5.4 若 $T \in \mathrm{Hom}(V, W)$, 试证 T(的图) 是 $V \times W$ 中 $W' = \{0\} \times W$ 的一个补.

5.5 若 l 是 V 上一个线性泛函 ($l \in \mathrm{Hom}(V, \mathbb{R}) = V^*$), 及若 α 是 V 中一个向量满足 $l(\alpha) \neq 0$, 试证 $V = N \bigoplus M$, 此处 N 是 l 的零空间及 $M = \mathbb{R}\alpha$ 为 α 的线性生成. 这个结果关于在 $\mathbb{R}^3$ 中的补能说明什么?

5.6 试证一个向量空间 V 的一个子空间 N 的任意补 M 都同构于商空间 V/N.

5.7 我们再假定每一个子空间都有一个补. 试证若 $T \in \operatorname{Hom} V$ 不是单射的, 则在 $\operatorname{Hom} V$ 中存在一个非零的 S 使 $T \circ S = 0$. 再证若 $T \in \operatorname{Hom} V$ 不是满射的, 则在 $\operatorname{Hom} V$ 中存在一个非零的 S 使 $S \circ T = 0$.

5.8 利用上面习题作为部分论证, 试证 $T \in \operatorname{Hom} V$ 为单射当且仅当 $T \circ S = 0 \Rightarrow S = 0$, 及 T 为满射的当且仅当 $S \circ T = 0 \Rightarrow S = 0$. 因此我们有一个形式的关于单射性与满射性的特征化, 其含义为它们不管 S 与 T 为变换, 但仅仅只管 S 与 T 作为一个代数中元素的代数性质.

5.9 命 M 与 N 为一个向量空间 V 的补子空间, 而命 X 为一个子空间满足 $X \cap N = \{0\}$. 试证存在一个由 X 至 M 的线性单射.[提示: 考虑沿 N, V 在 M 上的投影 P]. 试证一个子空间 N 的任意两个补是同构的, 这里需证上面的单射是满射的当且仅当 X 是 N 的一个补.

5.10 回到上面习题的第一点, 命 Y 为 $P[x]$ 在 M 中的一个补. 试证 $X \cap Y = \{0\}$ 及 $X \bigoplus Y$ 为 N 的一个补.

5.11 命 M 为 V 的一个真子空间, 又命 $\{\alpha_i : i \in I\}$ 为 V 中一个有限集合. 置 $L = L(\{\alpha_i\})$ 并假定 $M + L = V$. 试证存在一个子集合 $J \subset I$ 使 $\{\alpha_i : i \in J\}$ 生成 M 的一个补.[提示: 考虑一个最大可能的子集 J 使 $M \cap L(\{\alpha_i\}_J) = \{0\}$.]

5.12 给定 $T \in \operatorname{Hom}(V, W)$ 及 $S \in \operatorname{Hom}(W, X)$, 试证

(a) $S \circ T$ 是满射的 $\Leftrightarrow S$ 是满射的, 且 $R(T) + N(S) = W$;

(b) $S \circ T$ 是单射的 $\Leftrightarrow T$ 是单射的, 且 $R(T) \cap N(S) = \{0\}$;

(c) $S \circ T$ 是一个同构 $\Leftrightarrow S$ 是满射的, T 是单射的, 且 $W = R(T) \bigoplus N(S)$.

5.13 假定 V 的每一个子空间都有一个补, 试证 $T \in \operatorname{Hom} V$ 满足 $T^{\alpha} = 0$ 当且仅当 V 有一个直和分解 $V = M \bigoplus N$, 使在 N 上 $T = 0$ 及 $T[M] \subset N$.

5.14 其次假定 $T^3 = 0$, 但 $T^2 \neq 0$. 试证 V 可以被写成 $V = V_1 \bigoplus V_2 \bigoplus V_3$, 此处 $T[V_1] \subset V_2, T[V_2] \subset V_3$, 而在 V_3 上 $T = 0$.(再假定一个向量空间的任何子空间都有一个补).

5.15 我们现在假定 $T^n = 0$, 但 $T^{n-1} \neq 0$. 当 $i = 1, \cdots, n-1$ 时, 置 $N_i = T^i$ 的零空间, 及命 V_1 为 V 中 N_{n-1} 的一个补. 试首先证明

$$T[V_1] \cap N_{n-2} = \{0\}$$

及 $T[V_1] \subset N_{n-1}$. 将 $T[V_1]$ 延拓为 N_{n-1} 中 N_{n-2} 的一个补 V_2, 并证明按这个方式, 我们可以构造 $V_1, \cdots, V_n$ 使

$$V = \bigoplus_1^n V_i, \quad T[V_i] \subset V_{i+1}, i < n,$$

及

$$T[V_n] = \{0\}.$$

求解一个线性方程 数学中很多重要问题都有下面的一般形式. 给定一个线性算子 $T: V \to W$, 及给定一个 $\eta \in W$, 关于 $\xi \in V$, 求解方程 $T(\xi) = \eta$. 用我们的术语, 存在一个解的条件恰为 η 在 T 的值空间中的条件. 在特殊情况下, 这个条件可以被另外的多少有用的等价表述来给出. 假定我们知道如何去认识 $R(T)$, 在这种情况下, 我们可以使之成为新的值域, 并假定 T 是满射的. 仍然剩下确定什么是我们解方程的意思的问题. 贯穿问题的所有重要例子的普遍原则是计算 T 的一个右逆元之一个解程序, 即一个线性算子 $S: W \to V$ 使 $T \circ S = I_W$, 此处 I_W 为 W 上的单位. 因此, 对于每一个 $\eta \in W$, 一个解程序按这样一种方式给出一个解向量 $\xi \in V$, 即求解 ξ 随 η 线性地变化. 将这个作为我们求解的意思, 则我们有下述基本的重述.

定理 5.3 *命 T 为由向量空间 V 至向量空间 W 的一个满射线性映射, 又命 N 为它的零空间. 则一个子空间 M 是 N 的一个补当且仅当 T 对 M 的限制是一个由 M 至 W 的同构. 映射 $M \mapsto (T \restriction M)^{-1}$ 是由所有这种补子空间 M 的集合至 T 的所有线性右逆元之集合的一个双射映射.*

证明 这应该是清楚的, 即一个子空间 M 是 T 的一个线性右逆元 (一个满足 $T \circ S = I_W$ 的映射 S) 的值域当且仅当 $T \restriction M$ 是对 W 的一个同构, 在这种情况下, $S = (T \restriction M)^{-1}$. 严格地讲, 右逆元必须由 W 至 V, 所以必须是 $R = \iota_M \circ S$, 此处 ι_M 是由 M 至 V 的恒等单射. 则 $(R \circ T)^2 = R \circ (T \circ R) \circ T = R \circ I_W \circ T = R \circ T$, 及 $R \circ T$ 是一个投影, 其值域为 M 及其零空间为 N(因为 R 是单射的). 因此 $V = M \bigoplus N$. 反之, 若 $V = M \bigoplus N$, 则由于 $M \cap N = \{0\}$, 所以 $T \restriction M$ 是单射的, 由于 $M + N = V$ 可以导出 $W = T[V] = T[M + N] = T[M] + T[N] = T[M] + \{0\} = T[M]$, 所以 $T \restriction M$ 是满射的. □

***T* 的多项式** 这一小节的材料在我们研究常系数微分方程及一个对称矩阵可以对角化的证明中将要被用到. 在线性代数中, 对于任何处理矩阵标准型的方法, 它几乎都是基本的.

若 $p_1(t) = \sum_0^m a_i t^i$ 与 $p_2(t) = \sum_0^n b_j t^j$ 为两个多项式, 则它们的乘积是多项式

$$p(t) = p_1(t) p_2(t) = \sum_0^{m+n} c_k t^k,$$

此处 $c_k = \sum_{i+j=k} a_i b_j = \sum_{i=0}^k a_i b_{k-i}$. 现在命 T 为 $\mathrm{Hom}(V)$ 是任意固定元素, 及对于任意多项式 $q(t)$, 命 $q(T)$ 为将 t 换成 T 后所得之变换, 即若 $q(t) = \sum_0^l c_k t^k$,

则 $q(T)=\sum_0^l c_kT^k$, 此处当然 T^l 为 l 个因子的复合乘积 $T\circ T\circ\cdots\circ T$. 于是, 复合的双线性性表明 (定理 3.3), 若 $p(t)=p_1(t)p_2(t)$, 则 $p(T)=p_1(T)\circ p_2(T)$. 特别, 在复合之下, T 的任意两个多项式都是可相互交换的. 更简单些, 关于加法的交换律导出

$$\text{若 } p(t)=p_1(t)+p_2(t), \quad \text{则 } p(T)=p_1(T)+p_2(T).$$

由多项式代数到代数 $\mathrm{Hom}(V)$ 的映射 $p(t)\mapsto p(T)$ 因此保持加法, 乘法, 及 (显然地) 标量乘法, 即它保持一个代数的所有运算, 所以它是一个所谓的 (代数) 同态.

"同态" 这个字是一个一般的术语, 它描述同样类型的两个代数系统之间的一个映射 θ, 使 θ 保持系统的运算. 因此向量空间之间的一个同态就是一个线性变换, 而群之间的一个同态为一个保持一个群运算的映射. 一个可以见到, 但是并不典型的后者之例子为对数函数, 它是由正实数的乘法群至 $\mathbb{R}$ 的加法群的一个同态. 对数函数实际上是一个双射同态, 从而它是一个*群同构*.

如果这是一个代数课, 我们将证明除法程序与多项式的次数性质可以推出下面定理.(见习题 5.16~ 5.20.)

定理 5.4 *若 $p_1(t)$ 与 $p_2(t)$ 为互素多项式, 则存在多项式 $a_1(t)$ 与 $a_2(t)$ 使*

$$a_1(t)p_1(t)+a_2(t)p_2(t)=1.$$

我们关于互素的含义是除去常数外, 没有公因子. 在证明下面定理时, 我们将假定这个定理及它之前讨论的结果.

若 $T[M]\subset M$[即 $T\restriction M\in\mathrm{Hom}(M)$], 我们就称一个子空间 $M\subset V$ 在 $T\in\mathrm{Hom}(V)$ 之下是*不变的*.

定理 5.5 *命 T 为 $\mathrm{Hom}V$ 中任意变换, 又命 q 为任意多项式. 则 $q(T)$ 的零空间 N 在 T 之下是不变的, 及若 $q=q_1q_2$ 是 q 的任意互素因子分解与 N_1 和 N_2 分别为 $q_1(T)$ 与 $q_2(T)$ 的零空间, 则 $N=N_1\bigoplus N_2$.*

证明 由于 $T\circ q(T)=q(T)\circ T$, 所以若 $q(T)(\alpha)=0$, 则 $q(T)(T\alpha)=T(q(T)(\alpha))=0$, 从而 $T[N]\subset N$. 注意由于 $q(T)=q_1(T)\circ q_2(T)$, 所以任意 α 在 N_2 中亦在 N 中, 从而 $N_2\subset N$. 类似地, $N_1\subset N$. 所以我们可以将 V 换成 N 及 T 换成 $T\restriction N$; 因此我们可以假定 $T\in\mathrm{Hom}N$ 及 $q(T)=q_1(T)\circ q_2(T)=0$.

现在我们选取 a_1 与 a_2 使 $a_1q_1+a_2q_2=1$. 由于 $p\mapsto p(T)$ 为一个代数同态, 所以我们有

$$a_1(T)\circ q_1(T)+a_2(T)\circ q_2(T)=I.$$

置 $A_1=a_1(T)$ 等等, 则 $A_1\circ Q_1+A_2\circ Q_2=I$, $Q_1\circ Q_2=0$, 及所有算子 A_i, Q_i 是相互可交换的. 最后, 置 $P_i=A_i\circ Q_i=Q_i\circ A_i, i=1,2$. 则 $P_1+P_2=I$ 及 $P_1P_2=P_2P_1=0$. 因此 P_1 与 P_2 为投影, 及 N 为它们的值域之直和:$N=V_1\bigoplus V_2$. 因为每一个值域是另一个投影之零空间, 所以我们改写这个式子为 $N=N_1\bigoplus N_2$, 此处 $N_i=N(P_i)$. 剩下需要我们证明者为 $N(P_i)=N(Q_i)$. 首先注意由于 $Q_1\circ P_2=Q_1\circ Q_2\circ A_2=0$, 所以我们有 $Q_1=Q_1\circ I=Q_1\circ(P_1+P_2)=Q_1\circ P_1$. 所以两个恒等关系 $P_i=A_i\circ Q_i$ 与 $Q_i=Q_i\circ P_i$ 表明 P_i 与 Q_i 每一个的零空间都被另一个所包含, 从而它们相等, 定理证完. □

推论 5.4 *命 $p(t)=\prod_{i=1}^m p_i(t)$ 为多项式 $p(t)$ 分解为互素因子的一个分解式, 命 T 为 $\mathrm{Hom}(V)$ 中的一个元素, 又命 $N_i=N(p_i(T)), i=1,\cdots,m$ 与 $N=N(p(T))$. 则在 T 之下, N 及所有的 N_i 都是不变的, 且 $N=\bigoplus_{i=1}^m N_i$.*

证明 关于 m 用归纳法来证明. 定理就是 $m=2$ 之情况, 若我们置 $q=\prod_2^m p_i(t)$ 及 $M=N(q(T))$. 则由定理可知 $N=N_1\bigoplus M$ 及在 T 之下, N_1 与 M 都是不变的. 限制 T 于 M 上, 则由归纳法假定可见 $M=\bigoplus_{i=2}^m N_i$ 及当 $i=2,\cdots,m$ 时, N_i 在 T 之下是不变的. 由引理 5.3 的推论即推出我们的结果. □

习题

5.16 大概读者已经知道 (或者可以看得出) 一个多项式 P 的次数 $d(P)$ 适合下面的定律

$$d(P+Q)\leqslant\max\{d(P),d(Q)\}$$

$$d(P\cdot Q)=d(P)+d(Q),\quad 其中P与Q皆非零.$$

零多项式的次数未被定义 (它可以为 $-\infty$!). 试在 P 的次数上用归纳法以证明对于任意两个多项式 P 与 D, 其中 $D\neq 0$, 皆存在多项式 Q 与 R 使 $P=DQ+R$ 及 $d(R)\leqslant d(D)$ 或 $R=0$.[提示: 若 $d(P)<d(D)$, 我们可以取 Q 与 R 是什么? 若 $d(P)\geqslant d(D)$, 及若 P 与 D 的首项分别为 ax^n 与 bx^m, 其中 $n\geqslant m$, 则多项式

$$P'=P-\left(\frac{a}{b}\right)x^{n-m}D$$

的次数小于 $d(P)$, 所以由归纳法假定可知 $P'=DQ'+R'$, 现在可以完成证明了.]

5.17 假定上面的结果, 试证 R 与 Q 是由 P 与 D 惟一地决定的.(假定还有 $P = DQ' + R'$), 及由次数的性质来证明 $R' = R$ 及 $Q' = Q$). 这两个结果在一起, 构成了多项式的除法算式.

5.18 若 P 为任意多项式

$$P(x) = \sum_{0}^{n} a_i x^i,$$

又若 t 为任意数, 则当然 $P(t)$ 是数

$$\sum_{0}^{n} a_i x^i.$$

试用除法算式证明, 对于任意多项式 P 及任意数 t, 都存在一个多项式 Q 使

$$P(x) = (x - t)Q(x) + P(t),$$

所以 $P(x)$ 能够被 $x - t$ 整除当且仅当 $P(t) = 0$.

5.19 命 P 与 Q 为非零多项式. 选取多项式 A_0 与 B_0 使在所有形如 $AP + BQ$ 的多项式中, 多项式

$$D = A_0 P + B_0 Q$$

非零且有最小次数. 试证 D 为 P 与 Q 的公因子.(假定 D 不能整除 P, 则利用除法算式来得到一个关于 A_0 与 B_0 的选择的矛盾.)

5.20 命 P 与 Q 为非零互素多项式. 这就是说, 若 E 是 P 与 Q 的公共因子 ($P = EP', Q = EQ'$), 则 E 为一个常数. 试证存在多项式 A 与 B 使 $A(x)P(x) + B(x)Q(x) = 1$.(使用上面习题.)

5.21 在定理 5.5 的范围内, 试证 $q_2(T) = Q_2$ 在 N_1 上的限制是一个同构 (由 N_1 至 N_1).

5.22 V 上一个对合为一个适合 $T^2 = I$ 的映射 $T \in \mathrm{Hom}\, V$. 试证若 T 是一个对合, 则 V 是一个直和 $V = V_1 \bigoplus V_2$, 此处对于每一个 $\xi \in V_1$, 有 $T(\xi) = \xi$(在 V_1 上, $T = I$), 且对于每一个 $\xi \in V_2$, 有 $T(\xi) = -\xi$(在 V_2 上, $T = -I$).(应用定理 5.5.)

5.23 我们早先注意到 (在一个习题中), 若 φ 为任意由一个集合 A 至一个集合 B 的映射, 则 $f \mapsto f \circ \varphi$ 是由 $\mathbb{R}^B$ 至 $\mathbb{R}^A$ 的一个线性映射 T_φ. 试证若 $\psi : B \to C$, 则

$$T_{\psi \circ \varphi} = T_\varphi \circ T_\psi.$$

(这个应该是合成的可结合性的一个直接推论.)

5.24 命 A 为任意集合, 又命 $\varphi : A \to A$ 使对于每一个 a, $\varphi \circ \varphi(a) = a$, 则 $T_\varphi : f \mapsto f \circ \varphi$ 为在 $V = \mathbb{R}^A$ 上的一个对合 (由于 $T_{\varphi \circ \psi} = T_\psi \circ T_\varphi$). 试证 $\mathbb{R}^{\mathbb{R}}$ 作为偶函数子空间与奇函数子空间的直和分解是由这样一个映射 $\varphi : \mathbb{R} \to \mathbb{R}$ 定义的 $\mathbb{R}^{\mathbb{R}}$ 上的一个对合引起的.

5.25 命 V 为 $\mathbb{R}^{\mathbb{R}}$ 包含可微函数的子空间, 并假定在微分之下, V 是不变的 ($f \in V \Rightarrow Df \in V$). 再假定在 V 上, 线性算子 $D \in \operatorname{Hom} V$ 满足 $D^2 - 2D - 3I = 0$. 试证明 V 是两个子空间 M 与 N 的直和, 其中在 M 上, $D = 3I$, 而在 N 上, $D = -I$. 实际上, 由此推出 M 是一个向量的线性生成, 类似地, 对 N 亦然. 如果你能够, 试找出这两个函数. ($f' = 3f \Rightarrow f = ?$)

* **线性映射的区组分解** 给定 $\operatorname{Hom} V$ 中的 T 及一个直和分解 $V = \bigoplus_1^n V_i$, 与其对应的投影 $\{P_i\}_1^n$, 我们可以考虑映射 $T_{ij} = P_i \circ T \circ P_j$. 尽管 T_{ij} 是由 V 至 V, 我们仍然要考虑它是由 V_j 至 V_i(严格地讲, 在这种情况下, 它是什么?). 类似于一个矩阵, 我们图解地将这些 T_{ij} 画在一个矩形阵之中, 如下所示关于 $n = 2$ 之情形.

T_{11}	T_{12}
T_{21}	T_{22}

此处, 由于 $T = \sum_{i,j} T_{ij}$, 我们称有双重指标的系为 T 关于 V 的给定直和分解的区组分解.

更一般些, 若 $T \in \operatorname{Hom}(V, W)$ 及 W 亦有一个具有对应投影 $\{Q_i\}_1^m$ 的直和分解 $W = \bigoplus_{i=1}^m W_i$, 则由 $T_{ij} = Q_i \circ T \circ P_j$ 定义的系 $\{T_{ij}\}$ 及被画成一个 $m \times n$ 矩形阵为 T 关于两个直和分解的区组分解.

当Hom V 中的 T 关于 V 的一个特别直和分解有一个特别关系时, 对应的区组图形有这样的特征, 即用一种生动的方式展示这些特殊性质; 然后, 这将有助于我们更好地了解 T 的性质及更易于用它来计算.

例如, 若 $V = V_1 \bigoplus V_2$, 则在 T 之下, V_1 是不变的 (即 $T[V_1] \subset V_1$) 当且仅当区组图形是上三角的, 如下图所示.

T_{11}	T_{12}
0	T_{22}

其次, 假定 $T^2 = 0$. 命 V_1 为 T 之值域及假定 V_1 有一个补 V_2, 则读者应该清楚地看到对应的区组图形是

0	T_{12}
0	0

这个形式称为严格的上三角; 它是一个上三角及主对角线上为零. 反之, 若 T 有某种严格的上三角 2×2 区组图形, 则 $T^2 = 0$.

若 R 是一个合成乘积 $R = ST$, 则它的区组分量可以用 S 与 T 的某些项将它表示出来. 因此

$$R_{ik} = P_i R P_k = P_i S T P_k = P_i S \left(\sum_{j=1}^{n} P_j \right) T P_k = \sum_{j=1}^{n} S_{ij} T_{jk}.$$

在此我们用到恒等关系 $I = \sum_{j=1}^{n} P_j$ 与 $P_j = P_j^2$. 2×2 的情况可以被画出如下:

$S_{11}T_{11} + S_{12}T_{21}$	$S_{11}T_{12} + S_{12}T_{22}$
$S_{21}T_{11} + S_{22}T_{21}$	$S_{21}T_{12} + S_{22}T_{22}$

由此我们可以读出一个事实, 它将在以后对我们有用: 若 T 是 2×2 上三角的 $(T_{21} = 0)$, 及若 T_{ii} 作为由 V_i 至 V_i 的一个映射是可逆的, 此处 $i = 1, 2$, 则 T 是可逆的及其逆为

T_{11}^{-1}	$-T_{11}^{-1}T_{12}T_{22}^{-1}$
0	T_{22}^{-1}

我们求得这个解是由简单地置乘积图形等于

I_1	0
0	I_2

并求解之; 当然在我们手中有了这个图形时, 它就可以简单地被验证为正确的了.

习题

5.26 试证若 $T \in \operatorname{Hom} V$, $V = \bigoplus_1^n V_i$, 又若 $\{P_i\}_1^n$ 为对应的投影, 则变换 $T_{ij} = P_i \circ T \circ P_j$ 之和为 T.

5.27 若 S 与 T 在 $\operatorname{Hom} V$ 之中, 及 $\{S_{ij}\}, \{T_{ij}\}$ 为关于 V 的某直和分解的区组分量, 试证当 $j \neq l$ 时, $S_{ij} \circ T_{lk} = 0$.

5.28 试验证若 T 关于直和分解 $V = V_1 \bigoplus V_2$ 有一个上三角区组图形, 则 V_1 在 T 之下是不变的.

5.29 试验证若图形是严格的上三角的, 则 $T^2 = 0$.

5.30 试证若 $V = V_1 \bigoplus V_2 \bigoplus V_3$ 及 $T \in \operatorname{Hom} V$, 则子空间 V_i 在 T 之下都是不变的当且仅当关于 T 的区组图形为

$$\begin{bmatrix} T_{11} & 0 & 0 \\ 0 & T_{22} & 0 \\ 0 & 0 & T_{33} \end{bmatrix}.$$

再证 T 是可逆的当且仅当对每一个 i, T_{ii} 是可逆的 (作为 $\operatorname{Hom} V_i$ 的一个元素).

5.31 假定 T 有一个 2×2 上三角区组图形及作为 $\operatorname{Hom} V_i$ 的一个元素, T_{ii} 是可逆的 ($i = 1, 2$), 试用构筑一个 2×2 区组图形, 即一个关于 T 的图形与一个在课文中给出的作为 T 的逆的图形之积来验证 T 是可逆的.

5.32 假定 T 如前面习题所示, 试用考虑两个方程 $T \circ S = I$ 及 $S \circ T = I$ 的区组形式以证明 $S = T^{-1}$ 必须有给定的区组图形.

5.33 关于一个 3×3 区组图形的严格的上三角的意义应该是什么? 什么是 T 的对应性质? 试证 T 有这个性质当且仅当它有一个严格的上三角区组图形 (见习题 5.14.)

5.34 假定 $\operatorname{Hom} V$ 中的 T 满足 $T^n = 0$(但 $T^{n-1} \neq 0$). 试证 T 有一个严格的上三角 $n \times n$ 区组分解 (应用习题 5.15).

1.6 双线性性

双线性映射 双线性映射的概念对于了解线性代数是很重要的, 这是由于它是对偶性原则的向量形式 (0.10 节).

定义 6.1 若 U, V 与 W 为向量空间, 又若一个由 $U \times V$ 至 W 的映射

$$\omega : \prec \xi, \eta \succ \mapsto \omega(\xi, \eta)$$

满足: 当固定其中一个变量后, 它是另一个变量的线性映射, 则称这一映射是双线性的.

即若我们将 ξ 固定, 则 $\eta \mapsto \omega(\xi, \eta)$ 是线性的 [所以属于$\operatorname{Hom}(V, W)$]; 若我们将 η 固定, 则类似地 $\omega(\xi, \eta)$ 在 $\operatorname{Hom}(U, W)$ 中作为 ξ 的一个函数. 这与积向量空间 $U \times V$ 上的线性性不是同样的概念. 例如, $\prec x, y \succ \to x + y$ 是一个由 $\mathbb{R}^2$ 至 $\mathbb{R}$ 的线性映射, 但它不是双线性的. 若 y 固定, 则映射 $x \mapsto x + y$ 是仿射的 (过 y 之平移), 但它不是线性的, 除非 $y = 0$. 另一方面, $\prec x, y \succ \mapsto xy$ 是由 $\mathbb{R}^2$ 至 $\mathbb{R}$ 的一个双线性映射, 但它不是线性的. 若 y 固定, 则映射 $x \mapsto yx$ 是线性的. 但两个

有序数偶之和不映射至它们的像之和:

$$\prec x,y \succ + \prec u,v \succ = \prec x+u, y+v \succ \mapsto (x+u)(y+v),$$

它不是像之和 $xy+uv$. 类似地, 如我们在 §2 中所见, 标量乘积 $(\boldsymbol{x},\boldsymbol{y})=\sum_1^n x_i y_i$ 是由 $\mathbb{R}^n \times \mathbb{R}^n$ 至 $\mathbb{R}$ 的双线性映射.

双线性性的线性含义部分地由下面的定理来阐明.

定理 6.1 若 $\omega : U \times V \to W$ 是双线性的, 则由对偶性, ω 等价于由 U 至 $\mathrm{Hom}(V,W)$ 的一个线性映射, 亦等价于由 V 至 $\mathrm{Hom}(U,W)$ 的一个线性映射.

证明 对于每一个固定的 $\eta \in V$, 命 ω_η 为映射 $\xi \mapsto \omega(\xi,\eta)$. 即 $\omega_\eta(\xi) = \omega(\xi,\eta)$. 则由双线性的假设可知 $\omega_\eta \in \mathrm{Hom}(U,W)$. 所以映射 $\eta \mapsto \omega_\eta$ 是由 V 至 $\mathrm{Hom}(U,W)$, 而其线性性是由于当 ξ 固定时, ω 关于 η 的线性性:

$$\omega_{c\eta+d\zeta}(\xi) = \omega(\xi, c\eta+d\zeta) = c\omega(\xi,\eta) + d\omega(\xi,\zeta) = c\omega_\eta(\xi) + d\omega_\zeta(\xi),$$

所以

$$\omega_{c\eta+d\zeta} = c\omega_\eta + d\omega_\zeta.$$

类似地, 若我们由 $\omega^\xi(\eta) = \omega(\xi,\eta)$ 定义 ω^ξ, 则 $\xi \mapsto \omega^\xi$ 是一个由 U 至 $\mathrm{Hom}(V,W)$ 的线性映射. 反之, 若 $\varphi : U \to \mathrm{Hom}(V,W)$ 是线性的, 则由 $\omega(\xi,\eta) = \varphi(\xi)\eta$ 定义的函数 ω 是双线性的. 而且, $\omega^\xi = \varphi(\xi)$, 所以 φ 为映射 $\xi \mapsto \omega^\xi$. □

我们将看到双线性性是经常出现的. 有时由上面定理给出的重新解释给出了新的洞察; 在另一些时候, 它似乎是无用的.

例如, 合成映射 $\prec S,T \succ \mapsto S \circ T$ 是双线性的, 而且定理 3.3 的推论 (实际上是说一个固定的 T 从右方的合成是一个线性映射) 只是双线性性的一个显式语句的一部分. 但是线性映射 $\mapsto$ 由 T 的合成是一个复杂的事物, 除了 $W=\mathbb{R}$ 之情况外, 我们并不需要.

定理 6.2 映射 $\omega(\boldsymbol{x},\boldsymbol{\alpha}) = \sum_1^n x_i \alpha_i$ 是由 $\mathbb{R}^n \times V^n$ 至 V 的双线性映射. 所以映射 $\boldsymbol{\alpha} \to \omega_{\boldsymbol{\alpha}}$ 是一个由 V^n 至 $\mathrm{Hom}(\mathbb{R}^n, V)$ 的线性映射, 而实际上, 它是一个同构.

证明 对于一个固定的 $\boldsymbol{\alpha}$, ω 关于 $\boldsymbol{x}$ 的线性性已由定理 1.2 证明, 而对于一个固定的 $\boldsymbol{x}$, 它关于 $\boldsymbol{\alpha}$ 的线性性可以由同法得到. 所以由定理 6.1 可知 $\boldsymbol{\alpha} \mapsto \omega_{\boldsymbol{\alpha}}$ 是线性的. 它的双射性隐含于定理 1.2 之中. □

应该注意的为我们可以用任意有限指标集合 I 恰如特殊集合 $\overline{n}$, 并得出 $\omega(\boldsymbol{x},\boldsymbol{\alpha})=\sum_{i\in I}x_i\alpha_i$ 是由 $\mathbb{R}^I\times V^I$ 至 V 的双线性映射及 $\boldsymbol{\alpha}\mapsto\omega_{\boldsymbol{\alpha}}$ 为一个由 V^I 至 $\mathrm{Hom}(\mathbb{R}^I,V^I)$ 的同构. 同时注意, 按第 1 节的术语 $\omega_{\boldsymbol{\alpha}}=L_{\boldsymbol{\alpha}}$.

推论 6.1 由 $\mathbb{R}^n\times\mathbb{R}^n$ 至 $\mathbb{R}$ 的标量积 $(\boldsymbol{x},\boldsymbol{a})=\sum_1^n x_ia_i$ 是双线性的; 所以, $\boldsymbol{a}\mapsto\omega_{\boldsymbol{a}}=L_{\boldsymbol{a}}$ 是由 $\mathbb{R}^n$ 至 $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R})$ 的一个同构.

自然同构 我们常常发现两个向量空间彼此由这样一种方式联系着, 即它们之间的一个特殊的同构被选了出来. 这种现象很难用一般的术语来表述, 但易于用例子来描述.

对偶性是这种"自然"同构的一个源泉. 例如, 一个 $m\times n$ 矩阵 $\{t_{ij}\}$ 是两个变量 $\prec i,j\succ$ 的一个实值函数, 这样它就是笛卡儿空间 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 中的一个元素. 我们也可以将 $\{t_{ij}\}$ 看成是空间 $\mathbb{R}^m$ 中的 n 个列向量的序列. 这是对偶观点, 此处我们固定 j, 而对于每一个 j, 得到 i 的一个函数. 由这个观点, $\{t_{ij}\}$ 是 $(\mathbb{R}^{\overline{m}})^{\overline{n}}$ 的一个元素. $\mathbb{R}^{\overline{m}\times\overline{n}}$ 与 $(\mathbb{R}^{\overline{m}})^{\overline{n}}$ 之间的对应性显然是一个同构. 这是一个自然同构的例子.

其次, 我们审视一下观察笛卡儿 n 维空间自身的各种途径. 定义一个有序 n 元组 (联组) 的标准方式是用归纳法. 有序三元组 $\prec x,y,z\succ$ 被定义为一个有序对 $\prec\prec x,y\succ,z\succ$, 及有序 n 元组 $\prec x_1,\cdots,x_n\succ$ 被定义为 $\prec\prec x_1,\cdots,x_{n-1}\succ,x_n\succ$. 因此在置 $\mathbb{R}^1=\mathbb{R}$ 后, 我们归纳地定义 $\mathbb{R}^n=\mathbb{R}^{n-1}\times\mathbb{R}$.

有序 n 元组也可以被定义为 $\overline{n}=\{1,\cdots,n\}$ 上的函数, 此处在 i 取值 x_i, 则

$$\prec x_1,\cdots,x_n\succ=\{\prec 1,x_1\succ,\cdots,\prec n,x_n\succ\},$$

及笛卡儿 n 维空间为 $\mathbb{R}^{\overline{n}}=\mathbb{R}^{\{1,\cdots,n\}}$.

最后, 我们常常希望将笛卡儿 $(n+m)$ 维空间看成为笛卡儿 n 维空间与笛卡儿 m 维空间的笛卡儿乘积, 所以现在我们取

$$\prec x_1,\cdots,x_{n+m}\succ \quad 为 \quad \prec\prec x_1,\cdots,x_n\succ,\prec x_{n+1},\cdots,x_{n+m}\succ\succ$$

及 $\mathbb{R}^{n+m}$ 为 $\mathbb{R}^n\times\mathbb{R}^m$.

在此, 如果我们再次对相同的 n 元组的两种不同模型配对, 则我们得到一个关于笛卡儿 n 维空间的对应模型之间的显然自然同构.

最后, 由定理 3.6 与 3.7 给出的笛卡儿积空间的特征性质将推出自然同构. 定理 3.6 说明在一个共同定义域 V 上的线性映射的 n 元组 $\{T_i\}_1^n$ 等价于一个单个的 n 元组值映射 T, 此处对于所有 $\xi\in V$, $T(\xi)=\prec T_1(\xi),\cdots,T_n(\xi)\succ$. (这又

是对偶性! $T_i(\xi)$ 是两个变量 i 与 ξ 的一个函数 .) 不难看出, T 与 $\{T_i\}_1^n$ 的这种恒等是由 $\prod_i \mathrm{Hom}(V, W_i)$ 至 $\mathrm{Hom}(V, \prod_i W_i)$ 的一个同构.

类似地, 定理 3.7 将一个线性映射的 n 元组 $\{T_i\}_1^n$ 与一个共同值域 V 的 n 元组变量的单个线性映射 T 等同起来, 这个恒等关系是由 $\prod_1^n \mathrm{Hom}(W_i, V)$ 至 $\mathrm{Hom}(\prod_1^n W_i, V)$ 的一个自然同构.

两个向量空间之间的任何同构以一种暂时的方式将它们等同起来. 暂时我们设想向量表示同样的抽象空间, 仅仅只要我们面前有同构. 如果我们将它们之间的同构换为另一个不同的同构, 则我们得到一个新的暂时等同. 另一方面, 自然同构产生永久的等同性. 在更深的含义下, 我们设想序对事物为同样事物的两个方面. 因此我们设想一个矩阵 "为" 或者是一个行向量序列, 或者是一个列向量序列, 或者是两个整数指标的一个单个函数. 我们将在下一章第 3 节之末, 再对这个问题作最后的考察.

* 现在我们可以对线性组合公式的定理作终结考察了. 定律 S1~S3正好表述了标量积 $x\alpha$ 是双线性的. 更精确地, 它们表述了从 $\mathbb{R} \times W$ 至 W 的映射 $S: \prec x, \alpha \succ \mapsto x\alpha$ 是双线性的. 用定理 6.1 的语言为 $x\alpha = \omega_\alpha(x)$, 及由该定理我们得知映射 $\alpha \mapsto \omega_\alpha$ 是由 W 至 $\mathrm{Hom}(\mathbb{R}, W)$ 的一个同构.

W 与 $\mathrm{Hom}(\mathbb{R}, W)$ 之间的这个同构延拓为由 W^n 至 $(\mathrm{Hom}(\mathbb{R}, W))^n$ 的一个同构, 由笛卡儿第二乘积同构, 它又自然地同构于 $\mathrm{Hom}(\mathbb{R}^n, W)$. 因此 W^n 自然地同构于 $\mathrm{Hom}(\mathbb{R}^n, W)$; 映射为 $\boldsymbol{\alpha} \mapsto L_{\boldsymbol{\alpha}}$, 此处 $L_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \sum_1^n x_i\alpha_i$.

特别, $\mathbb{R}^n$ 自然地同构于 $\mathbb{R}^n$ 上所有线性泛函所成的空间 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R})$, 此处线性泛函是 n 元组 $\boldsymbol{a}$, 它对应于由 $\omega_{\boldsymbol{a}}(\boldsymbol{x}) = \sum_1^n a_i x_i$ 定义的泛函 $\omega_{\boldsymbol{a}}$.

同样, $(\mathbb{R}^m)^n$ 自然地同构于 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$. 由于 $\mathbb{R}^{\overline{m} \times \overline{n}}$ 自然地同构于 $(\mathbb{R}^m)^n$, 所以空间 $\mathbb{R}^{\overline{m} \times \overline{n}}$ 与 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 是自然同构的. 这仅为 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 中一个变换 T 至一个 $m \times n$ 矩阵 $\{t_{ij}\}$ 的自然结合.

第二章　有限维向量空间

若一个向量空间有一个有限生成集合, 则我们已经定义了它是有限维的; 本章, 我们将注意力集中于这种空间, 尽管这个限制对于我们的某些讨论是不必需的. 我们将看到我们可以给每一个有限维空间 V 一个惟一的整数, 它被称为 V 的维数. 关于维数它满足我们的直观需要, 而且它成为更深地探索这种空间性质的一个主要工具. 在这些进一步的研究中, "维数恒等" 的这个数字是决定性的. 我们将发现 V 上所有线性泛函的对偶空间 $V^* = \mathrm{Hom}(V, \mathbb{R})$ 在有限维理论中比一般向量空间的论述中起着更好的作用.(无论如何, 以本书的后面部分, 我们将会看到当我们将极限理论加入到代数中时, 存在一些无限维向量空间, 其中对偶空间亦起着同样重要的作用.) 一个有限维空间可以特征化为一个向量空间, 它同构于某笛卡儿空间 $\mathbb{R}^n$, 而且这样一个同构允许 $\mathrm{Hom}V$ 中一个变换 T 来 "转移" 至 $\mathbb{R}^n$, 于是它有一个矩阵. 所以这种空间上的线性变换理论完全地被矩阵理论所反映. 在本章中, 我们将比第一章所讲的更深入地进入这种关系的性质. 我们亦包括一节来讲矩阵计算; 一个简短节来讲述迹与行列式函数, 与一个二次型的对角化的短的讨论.

2.1　基

再考虑 V 中一个固定的有限有指标的向量集合$\alpha = \{\alpha_i : i \in I\}$, 及由 $\mathbb{R}^I$ 至

V 的对应线性组合映射 $L_\alpha : x \mapsto \sum x_i\alpha_i$, 它以 α 为骨架.

定义 1.1 若上面的映射 L_α 为单射, 则称有限有指标集合 $\{\alpha_i : i \in I\}$ 是独立的, 若 L_α 是一个同构 (到 V 上), 则称 $\{\alpha_i\}$ 为 V 的一个基.在这种情况下, 我们称 $\{\alpha_i : i \in I\}$ 为一个有序基或标架,此处 $I = \overline{n} = \{1, \cdots, n\}$, 其中 n 为某一正整数.

因此 $\{\alpha_i : i \in I\}$ 为一个基当且仅当对于每一个 $\xi \in V$, 皆存在一个惟一的有指标的 "系数" 集合 $x = \{x_i : i \in I\} \in \mathbb{R}^I$ 使 $\xi = \sum x_i\alpha_i$. 因为 $\{\alpha_i : i \in I\}$ 生成 V, 所以数 x_i 总是存在的. 又由于 L_α 是单射的, 所以 x 是惟一的.

例如, 我们可以直接地验证 $b^1 = \prec 2, 1 \succ$ 与 $b^2 = \prec 1, -3 \succ$ 构成 $\mathbb{R}^2$ 的一个基. 问题在于证明, 对于每一个 $y \in \mathbb{R}^2$, 皆存在一个惟一的 x 使

$$y = \sum_1^2 x_i b^i = x_1 \prec 2, 1 \succ + x_2 \prec 1, -3 \succ = \prec 2x_1 + x_2, x_1 - 3x_2 \succ .$$

由于这个向量方程等价于两个标量方程 $y_1 = 2x_1 + x_2$ 与 $y_2 = x_1 - 3x_2$, 用通常中学代数中的消去法, 我们即得惟一的解 $x_1 = (3y_1 + y_2)/7$ 与 $x_2 = (y_1 - 2y_2)/7$.

这些定义的形式是通过我们把线性组合公式解释为一个线性映射而引导出来的, 更常见的独立性定义是一个推论.

引理 1.1 有限有指标的集合 $\{\alpha_i : i \in I\}$ 的独立性等价于下面的性质, 即仅当所有系数 x_i 皆为 0 时才有 $\sum_I x_i\alpha_i = 0$.

证明 这个性质为 L_α 的零空间只包含 0, 所以它等价于 L_α 的单射性, 即由第一章引理 1.1 可知它等价于 $\{\alpha_i\}$ 的独立性. □

若 $\{\alpha_i\}_1^n$ 是 V 的一个有序基 (标架), 则满足 $\xi = \sum_1^n x_i\alpha_i$ 的惟一的 n 元组 x 被称为 ξ (关于基 $\{\alpha_i\}$) 的坐标 n 元组, x_i 称为 ξ 的第 i 个坐标. 我们称 $x_i\alpha_i$(有时为 x_i) 为 ξ 的第 i 个分量. 映射 L_α 称为一个基同构, 及它的逆 L_α^{-1} 称为一个坐标同构, L_α^{-1} 使每一个指定的向量 $\xi \in V$ 对应于其惟一的坐标 n 元组 x. 线性泛函 $\xi \to x_j$ 是第 j 个坐标泛函; 它是坐标同构 $\xi \mapsto x$ 与 $\mathbb{R}^n$ 上第 j 个坐标投影 $x \mapsto x_j$ 之复合. 我们将在 §3 看到 n 个坐标泛函形成 $V^* = \mathrm{Hom}(V, \mathbb{R})$ 的一个基。

在上一段中, 我们取指标集合 I 为 $\overline{n} = \{1, \cdots, n\}$, 并使用了语言 n 元组. 这与一个任意有限指标集合 I 的仅有区别为我们用一个坐标函数 $x = \{x_i : i \in I\}$

来代替一个坐标 n 元组.

我们首先关注的为将证明每一个有限维 (有限生成的) 向量空间都有一个基. 我们从指标的某些注记开始.

我们首先注意, 一个有限有指标的集合 $\{\alpha_i : i \in I\}$ 仅当指标作为一个至 V 的映射是单射的, 才可能是独立的, 这是由于若 $\alpha_k = \alpha_l$, 则 $\sum x_i\alpha_i = 0$, 此处 $x_k = 1, x_l = -1$, 及对于剩下的指标, $x_i = 0$. 同样, 若 $\{\alpha_i : i \in I\}$ 是独立的及 $J \subset I$, 则 $\{\alpha_i : i \in J\}$ 是独立的, 这是由于若 $\sum_J x_i\alpha_i = 0$, 及若我们置 $x_i = 0, i \in I - J$, 则 $\sum_I x_i\alpha_i = 0$, 所以每一个 x_j 都是 0. 一个有限无指标的集合当它在某 (必须是双射的) 指标下是独立的, 才称为独立的. 当然它将关于任何双射指标都是独立的. 一个任意集合, 若它的每一个有限子集合都是独立的, 则它就是独立的. 于是一个集合是相关的(不独立) 当且仅当存在 A 中不同的元素 $\alpha_1, \cdots, \alpha_n$ 及不全为零的标量 $x_1, \cdots, x_n$ 使 $\sum_1^n x_i\alpha_i = 0$. 一个无指标的基可以用显然的方式来定义. 无论如何, 一个集合总可以被认为是有指标的, 如果需要, 由其自身即可定义!

引理 1.2 若 B 为一个向量空间 V 的一个独立子集合, β 为任意不在线性生成 $L(B)$ 中的向量, 则 $B \cup \{\beta\}$ 是独立的.

证明 如果不然, 即存在一个零线性组合 $x\beta + \sum_1^n x_i\beta_i = 0$, 此处 $\beta_1, \cdots, \beta_n$ 为 B 的不同元素及系数不全为 0. 则 x 不能是零: 如果 $x = 0$, 则这一方程与 B 的独立性相矛盾, 所以我们可以除以 x 并解出 β, 结果是 $\beta \in L(B)$, 矛盾. □

读者将记住当一个向量空间 V 有一个有限生成集合 $\{\alpha_i\}_1^n$, 我们就称 V 是有限维的. 对于这样一个 V, 我们可以用上面引理选择一些 α_i 来构造 V 的一个基. 我们仅仅通过序列 $\{\alpha_i\}_1^n$, 并选择那些成员使增加以前选择的成员之线性生成. 由于 $\{\alpha_i\}_1^n$ 是生成集合, 所以我们将终止于一个生成集合. 由引理可知我们的子序列在每一步都是独立的. 用同样的方式, 我们可以在一个生成集合 $\{\alpha_i\}_1^n$ 中选取某些成员, 将一个独立集合 $\{\beta_i\}_1^m$ 拓展成一个基. 这个程序是直观的, 但将其严格地建立起来, 却很琐碎. 所以我们将另寻途径.

定理 1.1 任何最小的有限生成集合是一个基, 所以任何有限维向量空间有一个基. 更一般些, 若 $\{\beta_j : j :\in J\}$ 是有限独立集合及 $\{\alpha_i : i \in I\}$ 是一个有限生成集合, 又若 K 是 I 的一个最小子集合使 $\{\beta_j\}_J \cup \{\alpha_i\}_K$ 为生成集合, 则这个集合是独立的并为一个基, 所以一个有限维空间的任何有限独立子集合都可以被延拓为一个基.

证明 只要证明第二个论断即可, 这是由于它包含第一个论断作为特例. 若 $\{\beta_j\}_J \cup \{\alpha_i\}_K$ 不独立, 则存在一个非平凡的零线性组合 $\sum_J y_j\beta_j + \sum_K x_i\alpha_i = 0$. 若每一个 x_i 皆为零, 则这个方程将与 $\{\beta_j\}_J$ 的独立性相矛盾. 所以有某个 x_k 不是零, 从而我们可以从方程中解出 α_k. 即若我们置 $L = K - \{k\}$, 则 $\{\beta_j\}_J \cup \{\alpha_i\}_L$ 的线性生成包含 α_k. 所以它包含整个原来的生成集合, 从而它就是 V. 由于 L 是 K 的一个真子集合, 所以这与 K 的极小性质相矛盾. 因此 $\{\beta_j\}_J \cup \{\alpha_i\}_K$ 是独立的. □

其次, 我们注意到 $\mathbb{R}^n$ 自身有一个很特殊的基. 在指标映射 $i \mapsto \alpha_i$ 中, 向量 α_j 对应于指标 j, 但在线性组合映射 $\boldsymbol{x} \mapsto \sum x_i\alpha_i$ 下, 向量 α_j 对应于函数 δ^j, 它在 j 取值 1, 而其他则取值 0, 所以 $\sum_i \delta_i^j \alpha_i = \alpha_j$. 这个函数 δ^j 被称为*克罗内克 δ 函数*. 显然 1 - 点集合 $B = \{j\}$ 的特征函数 χ_B 与记号 'δ^j' 是含糊的, 正如 'χ_B' 是含糊的一样; 在每一种情况下, 含义依赖于隐含于课文中的定义域是什么. 我们已经在证明第一章定理 1.2 时用过了 $\mathbb{R}^n$ 上的 δ 函数.

定理 1.2 *克罗内克函数 $\{\delta^j\}_{j=1}^n$ 构成 $\mathbb{R}^n$ 的一个基.*

证明 由 δ^i 的定义可知 $\sum_1^n x_i\delta^i(j) = x_j$, 所以 $\sum_1^n x_i\delta^i$ 为 n 元组 $\boldsymbol{x}$ 自身, 从而线性组合映射 $L_\delta : \boldsymbol{x} \mapsto \sum_1^n x_i\delta^i$ 是恒等映射 $\boldsymbol{x} \mapsto \boldsymbol{x}$, 即一个平凡的同构. □

在 $\mathbb{R}^n$ 所有可能的有指标基中, 克罗内克基被拣出来是基于这样的事实, 即它的基同构是恒等映射; 基于这个理由, 它被称为 $\mathbb{R}^n$ 的标准基或自然基. 关于任意有限集合 I 对应的 $\mathbb{R}^I$ 亦成立.

最后, 我们将从一个基的存在性引出某些初等的结论.

定理 1.3 *若 $T \in \mathrm{Hom}(V, W)$ 为一个同构及 $\boldsymbol{\alpha} = \{\alpha_i : i \in I\}$ 为 V 的一个基, 则 $\{T(\alpha_i) : i \in I\}$ 为 W 的一个基.*

证明 由假定 $L_{\boldsymbol{\alpha}}$ 是 $\mathrm{Hom}(\mathbb{R}^n, V)$ 中一个同构, 所以 $T \circ L_{\boldsymbol{\alpha}}$ 为 $\mathrm{Hom}(\mathbb{R}^n, W)$ 中一个同构, 因此它的骨架 $\{T(\alpha_i)\}$ 为 W 的一个基. □

我们可以将任意基 $\{\alpha_i\}$ 看作在基同构之下, 标准基 $\{\delta^i\}$ 的像. 反之, 任意同构 $\theta : \mathbb{R}^I \to B$, 对于基 $\alpha_j = \theta(\delta^j)$ 变成一个基同构.

定理 1.4 *若 X 与 Y 为一个向量空间 V 的补子空间, 则 X 的一个基与 Y 的一个基的并为 V 的一个基. 反之, 若 V 的一个基分拆成两个集合, 它们分别*

线性生成 X 与 Y, 则 X 与 Y 为 V 的补子空间.

证明 我们仅证明第一个语句. 若 $\{\alpha_i : i \in J\}$ 是 X 的一个基及 $\{\alpha_i : i \in K\}$ 是 Y 的一个基, 则显然 $\{\alpha_i : i \in J \cup K\}$ 生成 V, 这是由于它的生成包括 X 与 Y, 所以 $X + Y = V$. 然后假定 $\sum_{J\cup K} x_i\alpha_i = 0$. 置 $\xi = \sum_J x_i\alpha_i$ 与 $\eta = \sum_K x_i\alpha_i$, 则我们得知 $\xi \in X, \eta \in Y$, 及 $\xi + \eta = 0$. 由于 X 与 Y 为互补, 所以 $\xi = \eta = 0$. 因此由于 $\{\alpha_i\}_J$ 是独立的, 所以 $x_i = 0(i \in J)$, 并因为 $\{\alpha_i\}_K$ 是独立的, 从而 $x_i = 0(i \in K)$. 所以, $\{\alpha_i\}_{J\cup K}$ 是 V 的一个基. 我们将逆部分之论证留作习题.□

推论 1.1 若 $V = \bigoplus_1^n V_i$ 及 B_i 为 V_i 的一个基, 则 $B = \bigcup_1^n B_i$ 为 V 的一个基.

证明 由上面定理可知 $B_1 \cup B_2$ 是 $V_1 \bigoplus V_2$ 的一个基. 由归纳法可知当 $j = 2, \cdots, n$ 时, $\bigcup_{i=1}^j B_i$ 是 $\bigoplus_{i=1}^j V_i$ 的一个基. 推论即 $j = n$ 的情况. □

如果我们将一个坐标同构接着一个线性组合映射, 则我们得到下面存在性定理的映射, 在此我们仅陈述 n 元组形式.

定理 1.5 若 $\beta = \{\beta_i\}_1^n$ 是向量空间 V 的一个有序基, 及若 $\{\alpha_i\}_1^n$ 为一个向量空间 W 中任意向量的 n 元组, 则存在一个惟一的 $S \in \mathrm{Hom}(V, W)$ 使 $S(\beta_i) = \alpha_i(i = 1, \cdots, n)$.

证明 由假定 $L_{\boldsymbol{\beta}}$ 是 $\mathrm{Hom}(\mathbb{R}^n, V)$ 中一个同构, 所以 $S = L_{\boldsymbol{\alpha}} \circ (L_{\boldsymbol{\beta}})^{-1}$ 为 $\mathrm{Hom}(V, W)$ 中的一个元素, 且满足 $S(\beta_i) = L_{\boldsymbol{\alpha}}(\delta^i) = \alpha_i$. 反之, 若 $S \in \mathrm{Hom}(V, W)$ 使对于所有 i, 皆有 $S(\beta_i) = \alpha_i$, 则 $S \circ L_{\boldsymbol{\beta}}(\delta^i) = \alpha_i$ 对所有 i 成立, 所以 $S \circ L_{\boldsymbol{\beta}} = L_{\boldsymbol{\alpha}}$. 因此 S 被惟一地确定为 $L_{\boldsymbol{\alpha}} \circ (L_{\boldsymbol{\beta}})^{-1}$. □

很自然地要询问, 上面惟一的 S 是如何随 n 元组 $\{\alpha_i\}$ 而变化的. 回答是: 线性性与"同构地".

定理 1.6 命 $\{\beta_i\}_1^n$ 是向量空间 V 的一个固定的有序基, 及对于向量空间 W 中选择的每一个 n 元组 $\boldsymbol{\alpha} = \{\alpha_i\}_1^n$, 命 $S_{\boldsymbol{\alpha}} \in \mathrm{Hom}(V, W)$ 为如上定义的惟一的变换. 则映射 $\boldsymbol{\alpha} \mapsto S_{\boldsymbol{\alpha}}$ 是由 W^n 至 $\mathrm{Hom}(V, W)$ 的一个同构.

证明 如上, $S_{\boldsymbol{\alpha}} = L_{\boldsymbol{\alpha}} \circ \theta^{-1}$, 此处 θ 为基同构 $L_{\boldsymbol{\beta}}$. 现在由第一章定理 6.2 我们得知 $\alpha \mapsto L_{\boldsymbol{\alpha}}$ 是一个由 W^n 至 $\mathrm{Hom}(\mathbb{R}^n, W)$ 的同构, 及由第一章定理 3.3 的推论可知, 一个固定的坐标同构 θ^{-1} 从右方合成则为由 $\mathrm{Hom}(\mathbb{R}^n, W)$ 至 $\mathrm{Hom}(V, W)$

的一个同构. 合成这两个同构即得定理. □

* **无穷基** 大部分向量空间都没有有限基, 所以很自然地, 我们试图将上述讨论推广至可能是无穷的指标集合 I. 克罗内克函数 $\{\delta^i : i \in I\}$ 有同样的定义, 但它们不再生成 $\mathbb{R}^I$. 由定义 f 是函数 δ^i 的一个线性组合当且仅当 f 有形式 $\sum_{i\in I_1} c_i\delta^i$, 此处 I_1 为 I 的一个有限子集合. 然后在 I_1 之外, 则 $f = 0$. 反之, 若除在一个有限集合 I_1 上之外, $f \in \mathbb{R}^I$ 皆取 0, 则 $f = \sum_{i\in I_1} f(i)\delta^i$. 所以 $\{\delta^i : i \in I\}$ 的线性生成正好是 $\mathbb{R}^I$ 上有下面性质的所有函数之集合, 即除去在一个有限集合上之外, 都是零. 我们记这个子空间为 $\mathbb{R}_I$.

若 $\{\alpha_i : i \in I\}$ 为 V 中向量的一个有指标集合及 $f \in \mathbb{R}_I$, 则若我们采用合理的规定, 即有任意个 0 的和为 0, 那么和 $\sum_{i\in I} f(i)\alpha_i$ 就变成有意思的. 因此, $\sum_{i\in I} = \sum_{i\in I_0}$, 此处 I_0 为 I 的任意有限子集合, 在它之外, f 为零.

有了这个规定, 如第一章定理 1.2, $L_\alpha : f \mapsto \sum_i f(i)\alpha_i$ 是由 $\mathbb{R}_I$ 至 V 的一个线性映射. 在同样的规定下, $\sum_{i\in I} f(i)\alpha_i$ 是关于向量 α_i 的一般线性组合的一个精致表达式. 代替去选择一个有限子集合 I_1 及正好对于那些 I_1 中的指标 i 对应的数 c_i, 我们对于所有 $i \in I$, 定义 c_i, 但具有规定, 即除去有限个指标外有 $c_i = 0$. 即我们取 $\boldsymbol{c} = \{c_i : i \in I\}$ 作为 $\mathbb{R}_I$ 中的一个函数.

如前, 我们给出关于*独立性*与*基*的同样定义, 则 $\{\alpha_i : i \in I\}$ 为 V 的一个基当且仅当 $L_\alpha : \mathbb{R}_I \to V$ 是一个同构, 即当且仅当对于每一个 $\xi \in V$ 皆存在一个惟一的 $\boldsymbol{x} \in \mathbb{R}_I$ 使 $\xi = \sum_i x_i\alpha_i$.

利用集合论公理的一条所谓的选择公理, 我们可以证明在这个含义下,每一个向量空间都有一个基及任意独立集合皆可以延拓为一个基。则仅对记号作一点改动即得定理 1.4 与 1.5. 特别, 若 V 的一个子集合 M 的一个基被延拓为 V 的一个基, 则增加部分的线性生成是补于 M 的一个子空间 N. 因此在纯代数的含义下, 每一个子空间都有一个补子空间. 在一些习题中, 我们将假定这个事实.

上面的和总是有限的 (不管外形), 而上述基的概念是纯代数的. 然而, 在这个意义下, 无穷基在分析中不是非常有用的, 所以我们目前将集中于有有限基的空间 (即有限维的). 然后在以后的一个重要课文中, 我们将讨论无穷基, 此处依靠极限理论, 和将是真正无穷的.

习题

1.1 试用直接计算证明 $\{\prec 1, -1 \succ, \prec 0, 1 \succ\}$ 是 $\mathbb{R}^2$ 的一个基.

1.2 学生必须了解一个向量的第 i 个坐标依赖于整个基而不只是第 i 个基向量. 试用标准基与上面习题的基就 $\mathbb{R}^2$ 中关于向量的第二个坐标证明这一结论.

1.3 试证 $\{\prec 1,1 \succ, \prec 1,2 \succ\}$ 是 $V = \mathbb{R}^2$ 的一个基. 这个由 $\mathbb{R}^2$ 至 V 的基同构现在是由 $\mathbb{R}^2$ 至 $\mathbb{R}^2$. 试找出坐标同构的矩阵.关于这个基, 试计算 $\prec -1,1 \succ, \prec 0,1 \succ$ 与 $\prec 2,3 \succ$ 的坐标.

1.4 试证 $\{\boldsymbol{b}^i\}_1^3$, 此处 $\boldsymbol{b}^1 = \prec 1,0,0 \succ, \boldsymbol{b}^2 = \prec 1,1,0 \succ$, 及 $\boldsymbol{b}^3 = \prec 1,1,1 \succ$ 是 $\mathbb{R}^3$ 的一个基.

1.5 在上面习题中, 试找出三个线性泛函 l_i, 它们是关于给予基的坐标泛函. 因为

$$\boldsymbol{x} = \sum_1^3 l_i(\boldsymbol{x})\boldsymbol{b}^i,$$

则寻找 l_i 就等价于求解 $\boldsymbol{x} = \sum_1^3 y_i \boldsymbol{b}^i$, 即用 $\boldsymbol{x} = \prec x_1, x_2, x_3 \succ$ 表示各 y_i.

1.6 试证任意没有两个多项式有同样的次数的多项式集合是独立的.

1.7 试证若 $\{\alpha_i\}_1^n$ 是 V 的一个独立子集合, $\mathrm{Hom}(V,W)$ 中的 T 是单射的, 则 $\{T(\alpha_i)\}_1^n$ 是 W 的一个独立子集合.

1.8 试证若 T 是 $\mathrm{Hom}(V,W)$ 中的任意元素, $\{T(\alpha_i)\}_1^n$ 在 W 中是独立的, 则 $\{\alpha_i\}_1^n$ 在 V 中是独立的.

1.9 若 V 的每一个基正好含有 n 个元素, 则在以后我们称向量空间 V 为 n 维的. 若 V 由一个单个向量 α 生成, 则 $V = \mathbb{R}\alpha$, 所以 V 显然是一维的.

命 $\{V_i\}_1^n$ 为一个向量空间 V 的一维子空间的一个集合, 及对于每一个 i, 在 V_i 中选取一个非零向量 α_i. 试证 $\{\alpha_i\}_1^n$ 是独立的当且仅当子空间 $\{V_i\}_1^n$ 是独立的, 而且 $\{\alpha_i\}_1^n$ 是一个基当且仅当 $V = \bigoplus_1^n V_i$.

1.10 试完成定理 1.4 的证明.

1.11 试基于同构的存在性给予定理 1.4 一个证明.

1.12 读者可能猜想, 及我们将在下一节证明, 一个有限维空间的每一个子空间都是有限维的. 现在往证一个有限维向量空间 V 的一个子空间 N 是有限维的当且仅当它有一个补 M.(由定理 1.1 与 1.4 及直和投影的联合来工作.)

1.13 由于 $\{\boldsymbol{b}^i\}_1^3 = \{\prec 1,0,0 \succ, \prec 1,1,0 \succ, \prec 1,1,1 \succ\}$ 是 $\mathbb{R}^3$ 的一个基, 所以在 $\mathrm{Hom}(\mathbb{R}^3,\mathbb{R}^2)$ 存在惟一的 T 使 $T(\boldsymbol{b}') = \prec 1,0 \succ, T(\boldsymbol{b}^2) = \prec 0,1 \succ$, 及 $T(\boldsymbol{b}^3) = \prec 1,1 \succ$. 试找出 T 的矩阵.(寻找 $T(\delta^i), i = 1,2,3$).

1.14 类似地, 试找出 $\mathrm{Hom}\mathbb{R}^3$ 中的 S 使 $S(\boldsymbol{b}^i) = \delta^i (i = 1,2,3)$.

1.15 试证明无穷序列 $\{t^n\}_0^\infty$ 是所有多项式的向量空间的一个基.

2.2 维数

维数的概念基于这样的事实, 即同样空间的两个不同的基包含同样多个元素. 因此这个数是 V 的每一个基的元素个数, 它被称为 V 的**维数**. 它给出了在同构下我们要了解的 V 的一切: 两个空间之间存在一个同构当且仅当它们有同样的维数. 我们将仅考虑有限维数. 如果 V 不是有限维的, 即它的维数是一个无穷基数, 这可能是一个读者不熟悉的概念.

引理 2.1 若 V 是有限维的及 $T \in \mathrm{Hom}V$ 是满射的, 则 T 是一个同构.

证明 命 n 为能生成 V 的最小元素个数. 即存在一个生成集合 $\{\alpha_i\}_1^n$, 但没有少于 n 个元素之生成集合. 则由定理 1.1 可知 $\{\alpha_i\}_1^n$ 是一个基, 从而线性组合映射 $\theta : \boldsymbol{x} \mapsto \sum_1^n x_i\alpha_i$ 是一个基同构. 由于 T 是满射的, 所以 $\{\beta_i\}_1^n = \{T(\alpha_i)\}_1^n$ 亦是生成集合, 从而 $T \circ \theta$ 亦是一个基同构. 因此, $T = (T \circ \theta) \circ \theta^{-1}$ 是一个同构. □

定理 2.1 若 V 是有限维的, 则 V 的所有基都有同样数目的元素.

证明 两个分别有 n 与 m 个元素的基决定了基同构 $\theta : \mathbb{R}^n \to V$ 与 $\varphi : \mathbb{R}^m \to V$. 假定 $m < n$, 及将 $\mathbb{R}^n$ 看成 $\mathbb{R}^m \times \mathbb{R}^{n-m}$, 命 π 为 $\mathbb{R}^n$ 在 $\mathbb{R}^m$ 上的投影.

$$\pi(\prec x_1, \cdots, x_m, \cdots, x_n \succ) = \prec x_1, \cdots, x_m \succ .$$

因为 $T = \theta^{-1} \circ \varphi$ 是一个由 $\mathbb{R}^m$ 至 $\mathbb{R}^n$ 的同构, 所以 $T \circ \pi : \mathbb{R}^n \to \mathbb{R}^n$ 是满射的, 因此由引理可知 $T \circ \pi$ 是一个同构. 则 $\pi = T^{-1} \circ (T \circ \pi)$ 是一个同构. 但因为 $\pi(\delta^n) = 0$, 这不可能, 因此我们得到一个矛盾. 所以没有一个基, 其元素个数可以少于任何其他基的元素个数. □

V 的每个基的元素个数这个整数当然被称为 V 的**维数**, 我们将它记为 $d(V)$. 因为 $\mathbb{R}^n$ 的标准基 $\{\delta^i\}_1^n$ 有 n 个元素, 所以在这个精确意义之下, $\mathbb{R}^n$ 是 n 维的.

推论 2.1 两个有限维向量空间是同构的当且仅当它们有同样的维数.

证明 若 T 是 V 至 W 的一个同构及 B 是 V 的一个基, 则由定理 1.3 可知 $T[B]$ 是 W 的一个基. 所以 $d(V) = \#B = \#T[B] = d(W)$, 此处 $\#A$ 表示 A 的元素个数. 反之, 若 $d(V) = d(W) = n$, 则 V 与 W 皆同构于 $\mathbb{R}^n$, 故它们彼此同构. □

定理 2.2 一个有限维向量空间 V 的每一个子空间 M 都是有限维的.

证明 命 $\mathfrak{a}$ 为 M 的有限独立子集合所成之族. 由定理 1.1 可知若 $A \in \mathfrak{a}$, 则 A 可以被延拓为 V 的一个基, 所以 $\#A \leqslant d(V)$. 因此 $\{\#A : A \in \mathfrak{a}\}$是一个整数的有限集合, 及我们可以选取 $B \in \mathfrak{a}$ 使 $n = \#B$ 是这个有限集合中的最大者. 于是 $L(B) = M$, 这是由于不然的话, 对于任何 $\alpha \in M - L(B)$, 由引理 1.2 我们有 $B \cup \{\alpha\} \in \mathfrak{a}$, 及

$$\#(B \cup \{\alpha\}) = n + 1,$$

这与 n 的最大性质相矛盾, 因此 M 是有限地生成的. □

推论 2.2 一个有限维空间 V 的每一个子空间 M 都有一个补.

证明 利用定理 1.1 将 M 的一个基延拓为 V 的一个基, 及命 N 为增加向量的线性生成. 然后应用定理 1.4. □

维数的恒等式 我们现在证明两个基本的维数的恒等式. 我们将总是假定 V 是有限维的.

引理 2.2 若 V_1 与 V_2 是 V 的互补子空间, 则 $d(V) = d(V_1) + d(V_2)$. 更一般些, 若 $V = \bigoplus_1^n V_i$, 则 $d(V) = \sum_1^n d(V_i)$.

证明 由定理 1.4 及其推论立即推出. □

定理 2.3 若 U 与 W 为一个有限维实向量空间的子空间, 则 $d(U + W) + d(U \cap W) = d(U) + d(W)$.

证明 命 V 为 $U \cap W$ 在 U 中的一个补. 我们开始证明 V 亦是 W 在 $U + W$ 中的一个补. 首先

$$V + W = V + ((U \cap W) + W) = (V + (U \cap W)) + W = U + W.$$

在此我们用到了明显的事实, 即一个向量空间与一个子空间的和仍为向量空间. 其次,

$$V \cap W = (V \cap U) \cap W = V \cap (U \cap W) = \{0\},$$

这是由于 V 是 $U \cap W$ 在 U 中的一个补. 因此我们有 $V + W = U + W$ 及 $V \cap W = \{0\}$, 所以由第一章引理 5.2 的推论可知 V 是 W 在 $U + W$ 中的一个补.

现在定理就是上面引理的一个推论. 我们有

$$\begin{aligned} d(U)+d(W) &= (d(U\cap W)+d(V))+d(W) \\ &= d(U\cap W)+(d(V)+d(W)) \\ &= d(U\cap W)+d(U+W). \end{aligned}$$ □

定理 2.4 命 V 为有限维的, 又命 W 为任意向量空间. 命 $T\in \mathrm{Hom}(V,W)$ 有零空间 N(在 V 中) 与值域 R(在 W 中), 则 R 是有限维的且 $d(V)=d(N)+d(R)$.

证明 命 U 为 N 在 V 中的一个补. 则我们知道 $T\restriction U$ 是至 R 的一个同构 (见第一章定理 5.3). 所以, R 是有限维的, 由我们的第一个恒等式可知 $d(R)+d(N)=d(U)+d(N)=d(V)$. □

推论 2.3 若 W 是有限维的及 $d(W)=d(V)$, 则 T 是单射的当且仅当它是满射的, 所以在这种情况下, 单射性、满射性与双射性都是等价的.

证明 T 是满射的当且仅当 $R=W$. 但这是等价于 $d(R)=d(W)$, 及若 $d(W)=d(V)$, 则定理表示为等价于 $d(N)=0$, 即等价于 $N=\{0\}$. □

定理 2.5 若 $d(V)=n$ 及 $d(W)=m$, 则 $\mathrm{Hom}(V,W)$ 为有限维的及其维数为 mn.

证明 由定理 1.6 可知 $\mathrm{Hom}(V,W)$ 同构于 W^n, 它在单射 $\theta_i(i=1,\cdots,n)$ 之下, 同构于 W 的 n 个子空间的直和. 所以由引理 2.2 可知 W^n 的维数为 $\sum_1^n m=mn$. □

定理 2.5 的另一证明将在第 4 节给出.

习题

2.1 试证若 $d(V)=n$, 则任何 n 个元素的生成子集合为一个基.

2.2 试证若 $d(V)=n$, 则任何 n 个元素的独立子集合为一个基.

2.3 试证若 $d(V)=n$ 及 W 为有同样维数的一个子空间, 则 $W=V$.

2.4 试用维数的恒等式证明, 若 f 是一个 n 维空间 V 上的一个非零线性泛函, 则其零空间有维数 $n-1$.

2.5 试用维数恒等式证明, 若 f 是一个有限维空间 V 上的一个线性泛函, 及 α 为不在它的零空间 N 中的一个向量, 则 $V = N \bigoplus \mathbb{R}\alpha$.

2.6 给定一个 n 维向量空间 V 的一个 $(n-1)$ 维子空间 N, 试证 N 是一个线性泛函的零空间.

2.7 命 X 与 Y 是一个有限维向量空间 V 的子空间, 及假定 $T \in \mathrm{Hom}(V, W)$ 有零空间 $N = X \cap Y$. 试证 $T[X+Y] = T[X] \bigoplus T[Y]$, 然后由引理 2.2 及定理 2.4 推出定理 2.3. 这个证明仍依赖于以 $N = X \cap Y$ 为其零空间的一个 T 的存在性. 我们知道不知道任意这样的 T?

2.8 试证若 V 是有限维的及 S, $T \in \mathrm{Hom}V$, 则

$$S \circ T = I \Rightarrow T\text{是可逆的}.$$

再证明 $T \circ S = I \Rightarrow T$ 是可逆的.

2.9 若商空间 V/N 是有限维的, 且有维数 n, 则称一个向量空间 V 的一个子空间 N 有有限余维数n. 试证一个子空间 N 有有限余维数 n 当且仅当 N 有一个维数 n 的补子空间 M(将 V/N 的一个基移回 V).不要假定 V 是有限维的.

2.10 试证若 N_1 与 N_2 为一个向量空间 V 的有有限维数的子空间, 则 $N = N_1 \cap N_2$ 有有限余维数及

$$\mathrm{cod}(N) \leqslant \mathrm{cod}(N_1) + \mathrm{cod}(N_2).$$

(当 $\overline{\xi}_i$ 为 N_i 包含 ξ 的陪集时, 考虑映射 $\xi \mapsto \prec \overline{\xi}_1, \overline{\xi}_2 \succ$.)

2.11 在上面习题中, 假定 $\mathrm{cod}(N_1) = \mathrm{cod}(N_2)$, 即 $d(V/N_1) = d(V/N_2)$. 试证 $d(N_1/N) = d(N_2/N)$.

2.12 给定 V 中非零向量 β 及 V^* 中 f 使 $f(B) \neq 0$, 试证映射 $\xi \mapsto f(\xi)\beta$ 的某标量积是一个投影. 再证任何投影都有一个按这一方式引起的一维值域.

2.13 我们知道在欧氏三维空间 $\mathbb{E}^3$ 中一个原点 O 的选择导致 $\mathbb{E}^3$ 中一个向量空间结构 (在对应 $X \mapsto \overrightarrow{OX}$ 之下), 及这个向量空间是三维的. 试证过 O 的一张几何平面变成一个二维子空间.

2.14 一个 m 维平面 M 是一个 m 维子空间 N 的一个平移 $N + \alpha_0$. 命 $\{\beta_i\}_1^n$ 为 N 的任意基, 及置 $\alpha_i = \beta_i + \alpha_0$. 试证 M 恰好是线性组合

$$\sum_{i=0}^{m} x_i\alpha_i, \quad \text{其中} \sum_{i=0}^{m} x_i = 1$$

所成之集合.

2.15 试证习题 2.14 是第一章习题 4.14 的一个推论.

2.16 反之, 试证若一张平面 M 是 $m+1$ 个元素的仿射生成, 则它的维数 $\leqslant m$.

2.17 由上面两个习题, 试编造一个仿射子空间的维数的一个直接定义.

2.18 试由下面的定义建议写一篇小论文. 若条件

$$\sum_0^m x_i\alpha_i = 0 \quad 与 \quad \sum_0^m x_i = 0$$

一起可以推出

$$x_i = 0 \quad 对于所有\ i\ 成立,$$

则称一个 $(m+1)$ 元组 $\{\alpha_i\}_0^m$ 是*仿射独立的*.

2.19 一个向量空间 V 上的一个多项式是 V 上的一个实值函数, 它可以被表示为线性泛函的有限乘积的一个有限和. 定义一个多项式的*次数*; 定义一个 *k 次齐次多项式*. 试证明 k 次齐次多项式的集合是一个向量空间 X_k.

2.20 继续上面的习题, 试证若 $k_1 < k_2 < \cdots < k_N$, 则向量空间 $\{X_{k_i}\}_1^N$ 为所有多项式之向量空间的独立子空间. [假定一个实变量的多项式 $p(t)$ 仅当它的所有系数皆为 0 时才是零多项式. 对于 V 上任意多项式 P, 考虑多项式 $p_\alpha(t) = P(t\alpha)$.]

2.21 命 $\prec \alpha, \beta \succ$ 为二维空间 V 的一个基, 及命 $\prec \lambda, \mu \succ$ 为对应的坐标投影 (V^* 的对偶基).试证每一个 V 上的多项式 "是两个变量 λ 与 μ 的多项式".

2.22 命 $\prec \alpha, \beta \succ$ 为一个二维向量空间 V 的一个基, 又命 $\prec \lambda, \mu \succ$ 为对应的坐标投影 (V^* 的对偶基). 试证

$$\prec \lambda^2, \lambda\mu, \mu^2 \succ$$

为 V 上次数为 2 的齐次多项式所成的向量空间之一个基. 类似地, 试计算二维向量空间上次数为 3 的齐次多项式空间之维数.

2.23 命 V 与 W 为二维向量空间, 及 F 为由 V 至 W 的映射. 试利用坐标系定义 F 为平方的概念, 然后证明它关于坐标系是独立的. 试将上面习题推广至较高维数及较高次数.

2.24 现在命 $F: V \to W$ 为二维空间之一个映射, 满足对于任意 $\boldsymbol{u}, \boldsymbol{v} \in V$ 及任意 $l \in W^*$, $l(F(t\boldsymbol{u}+\boldsymbol{v}))$ 为一个 t 的二次函数, 即其有形式 at^2+bt+c. 试证 F 是按照你在上面习题中定义的平方的.

2.3 对偶空间

尽管在这一节中, 所有空间都将被假定为有限维的. 但许多定义与性质对无限维空间亦正确. 在纯代数的情况与代数被混有连续性假设的情况之间是有一些区别的. 有限维数的幸运之一为没有这种复杂性. 正如读者可能从我们已经遇

到过的特殊线性泛函，特别是坐标泛函的个数，推测到 V 上所有线性泛函之空间 $\mathrm{Hom}(V,\mathbb{R})$ 将起着一个特殊的作用.

定义 3.1 向量空间 V 的对偶空间(或共轭空间)V^* 是所有从 V 至 $\mathbb{R}$ 的线性映射的向量空间 $\mathrm{Hom}(V,\mathbb{R})$，它的元素皆被称为线性泛函.

从某种意义上讲，我们将看到 V 亦是 V^* 的对偶空间 (V 与 $(V^*)^*$ 是自然地同构的)，所以两个空间是对称地相关联的. 我们将简单地研究一下零化(正交性)这个概念，它的源泉在于这种关联，然后我们将看到 $\mathrm{Hom}(V,W)$ 与 $\mathrm{Hom}(W^*,V^*)$ 之间有一个自然同构. 这给予数学家一个用来研究 $\mathrm{Hom}(V,W)$ 中的线性变换 T 的新的工具；T 与其像 T^* 之间的关系揭示了 T 本身的新性质.

对偶基 在开始时，人们自然地要问一个空间 V^* 有多大，我们立即来解决这个问题.

定理 3.1 命 $\{\beta_i\}_1^n$ 为 V 的一个有序基，又命 ε_j 为对应的 V 上第 j 个坐标泛函：$\varepsilon_j(\xi)=x_j$，此处 $\xi=\sum_1^n x_i\beta_i$. 则 $\{\varepsilon_j\}_1^n$ 是 V^* 的一个有序基.

证明 我们给出一个直接初等计算的证明.

(a) 独立性. 假定 $\sum_1^n c_j\varepsilon_j=0$，即对于所有 $\xi\in V,\sum_1^n c_j\varepsilon_j(\xi)=0$. 取 $\xi=\beta_i$ 并回忆一下，β_i 的坐标 n 元组为 δ^i，我们看到上面的方程化成了 $c_i=0$，这对所有 i 皆成立. 因此，$\{\varepsilon_j\}_1^n$ 是独立的.

(b) 生成性. 首先注意基展开 $\xi=\sum x_i\beta_i$ 可以被改写为 $\xi=\sum\varepsilon(\xi)\beta_i$. 则对于任意 $\lambda\in V^*$，我们有 $\lambda(\xi)=\sum_1^n l_i\varepsilon_i(\xi)$，此处我们置 $l_i=\lambda(\beta_i)$. 即 $\lambda=\sum l_i\varepsilon_i$. 这证明了 $\{\varepsilon_j\}_1^n$ 生成 V^*，及与 (a) 一起可知它是一个基. □

定义 3.2 V^* 的基 $\{\varepsilon_j\}$ 被称为 V 的基 $\{\beta_i\}$ 之对偶.

如通常的那样，我们的基本同构之一皆潜伏于这个之后，但我们将它的揭示留作一个习题.

推论 3.1 $d(V^*)=d(V)$.

三个方程

$$\xi=\sum\varepsilon_i(\xi)\beta_i,\quad \lambda=\sum\lambda(\beta_i)\varepsilon_i,\quad \lambda(\xi)=\sum\lambda(\beta_i)\cdot\varepsilon_i(\xi)$$

值得去考虑. 前面两个对称地关联着，每一个表示一个向量的基展开，其系数由

应用向量的对偶基的对应元素来计算. 第三个是 ξ 与 λ 之间的自身对称.

由于一个有限维空间 V 与它的对偶空间 V^* 有同样的维数, 所以, 当然它们是同构的. 事实上, 由于我们有由 V 至 $\mathbb{R}^n$ 相关联的坐标同构, 所以 V 的每一个基定义了一个同构, 即由 $\mathbb{R}^n$ 至 V^* 的对偶基同构, 从而有由 V 至 V^* 的复合同构. 这个同构随基而变化. 然而, 一般说来, 在 V 与 V^* 之间没有自然同构.

对于笛卡儿空间 $\mathbb{R}^n$ 则是另外一回事, 这是由于它有一个标准基, 所以与它的对偶空间 $(\mathbb{R}^n)^*$ 有一个标准同构. 不难看出, 这是同构 $\boldsymbol{a} \mapsto L_{\boldsymbol{a}}$, 此处 $L_{\boldsymbol{a}}(\boldsymbol{x}) = \sum_1^n a_i x_i$, 我们在 1.6 节中已经讨论过. 所以可以自由地去将 $\mathbb{R}^n$ 与 $(\mathbb{R}^n)^*$ 等同起来, 在此仅需铭记住的为当我们设想一个 n 元组 $\boldsymbol{a}$ 为一个线性泛函时, 我们的意思是泛函 $L_{\boldsymbol{a}}(\boldsymbol{x}) = \sum_1^n a_i x_i$.

第二共轭空间 一般说来, 尽管 V 与 V^* 不是自然地同构的, 我们现在将看到 V 自然地同构于 $V^{**} = (V^*)^*$.

定理 3.2 *由 $\omega(\xi, f) = f(\xi)$ 定义的函数 $\omega : V \times V^* \to \mathbb{R}$ 是双线性的, 及由 V 至 V^{**} 的映射 $\xi \mapsto \omega^\xi$ 是一个自然同构.*

证明 在这个课文中, 我们经常置 $\xi^{**} = \omega^\xi$, 所以 ξ^{**} 由对于所有 $f \in V^*$, $\xi^{**}(f) = f(\xi)$ 所定义. ω 的双线性性应该是清楚的, 所以第一章的定理 6.1 可以应用. 读者可由 $(c_1\xi_1 + c_2\xi_2)^{**}(f)$ 开始, 对 $\xi \mapsto \xi^{**}$ 的线性性作一个直接的验证.

这个映射的单射性仍然是问题. 若 $\alpha \neq 0$, 我们可以找到 $f \in V^*$ 使 $f(\alpha) \neq 0$; 一个方法使 α 为一个有序基的第一个向量并取 f 为对偶基中第一个泛函; 则 $f(\alpha) = 1$. 因为 $\alpha^{**}(f) = f(\alpha) \neq 0$, 我们特别看到 $\alpha^{**} \neq 0$. 所以映射 $\xi \to \xi^{**}$ 是单射, 并由定理 2.4 的推论可知它是双射的. □

如果我们设想 V^{**} 以这种方式自然地等同于 V, 则空间 V 与 V^* 对称地互相关联着. 每一个都是另一个的对偶. 在表达式 '$f(\xi)$' 中, 我们设想两个记号都是变量, 然后使一个或另一个固定从而有两种解释. 在这种情况下, 我们常常用一个更为对称的记号, 如 $\prec \xi, f \succ$, 来表示我们的意图, 即将两个记号都作为变量来对待.

引理 3.1 *若 $\{\lambda_i\}$ 是 V^* 的基, 它对偶于 V 的基 $\{\alpha_i\}$, 则 $\{\alpha_i^{**}\}$ 为 V^{**} 的基, 它对偶于 V^* 的基 $\{\lambda_i\}$.*

证明 我们有 $\alpha_i^{**}(\lambda_j) = \lambda_j(\alpha_i) = \delta_j^i$, 这表示 α_i^{**} 是第 i 个坐标投影. 在读者已忘记的情况下, 基表达式 $f = \sum c_j \lambda_j$ 可导出 $\alpha_i^{**}(f) = f(\alpha_i) = (\sum c_j \lambda_j)(\alpha_i) =$

c_i, 所以 α_i^{**} 是映射 $f \mapsto c_i$. □

零化子子空间 在这个对偶情况下, 正交性第一次自然地出现. 然而, 在以后的课文中, V 与 V^* 已经被一个标量积等同起来, 我们将省去 '正交' 这个词. 我们将在此说一个集合的零化子来代替它的正交补.

定义 3.3 若 $A \subset V$, 则 A 的零化子A° 是 V^* 中所有满足下面条件的 f 之集合: 对于所有 $\alpha \in A$ 有 $f(\alpha) = 0$.

类似地, 若 $A \subset V^*$, 则

$$A^\circ = \{\alpha \in V : 对于所有 f \in A, f(\alpha) = 0\}.$$

若我们将 V 看成 $(V^*)^*$, 则第二个定义就含于第一个之中.

下面的性质易于证明, 我们将它们留作习题:

(1) A° 总是一个子空间.

(2) $A \subset B \Rightarrow B^\circ \subset A^\circ$.

(3) $(L(A))^\circ = A^\circ$.

(4) $(A \cup B)^\circ = A^\circ \cap B^\circ$.

(5) $A \subset A^{\circ\circ}$.

我们现在对上一节那些维数的恒等式, 再增加较关键性的一个.

定理 3.3 若 W 是 V 的一个子空间, 则 $d(V) = d(W) + d(W^\circ)$.

证明 命 $\{\beta_i\}_1^m$ 为 W 的一个基, 及将它延拓为 V 的一个基 $\{\beta_i\}_1^n$. 命 $\{\lambda_i\}_1^n$ 为 V^* 中的对偶基. 我们宣称 $\{\lambda_i\}_{m+1}^n$ 为 W° 的一个基. 首先, 若 $j > m$, 则当 $i = 1, \cdots, m$ 时, $\lambda_j(\beta_i) = 0$, 所以由上面的 (3) 可知 λ_j 在 W° 之中. 因此 $\{\lambda_{m+1}, \cdots, \lambda_n\} \subset W^\circ$. 现在假定 $f \in W^\circ$, 及命 $f = \sum_{j=1}^n c_j\lambda_j$ 为它的 (对偶) 基展开式. 则对于每一个 $i \leqslant m$, 由于 $\beta_i \in W$ 及 $f \in W^\circ$, 所以我们有 $c_i = f(\beta_i) = 0$; 从而 $f = \sum_{m+1}^n c_j\lambda_j$. 因此 W° 中每一个 f 皆在 $\{\lambda_i\}_{m+1}^n$ 的生成之中. 综合上述, 我们已证明如我们宣称的 W° 是 $\{\lambda_i\}_{m+1}^n$ 的生成. 所以 $d(W^\circ) + d(W) = (n - m) + m = n = d(V)$. 定理证完. □

推论 3.2 对于每一个子集合 $A \subset V$ 有 $A^{\circ\circ} = L(A)$.

证明 由于 $(L(A))^\circ = A^\circ$, 所以由定理可知 $d(L(A)) + d(A^\circ) = d(V)$. 我们亦有 $d(A^\circ) + d(A^{\circ\circ}) = d(V^*) = d(V)$. 因此 $d(A^{\circ\circ}) = d(L(A))$, 及由上面 (5) 可知

$L(A) \subset A^{\circ\circ}$, 所以 $L(A) = A^{\circ\circ}$. □

T 的伴随 我们现在将看到, 对于 $\mathrm{Hom}(V, W)$ 中每一个 T, 皆存在 $\mathrm{Hom}(W^*, V^*)$ 与之自然地相关联的元素, 我们称它为 T 的伴随, 并记为 T^*. T 与 T^* 之间密切的关系的推论之一为 T^* 的值域恰好是 T 的零空间的零化子. 结合我们的维数恒等式可知这可以推出 T 与 T^* 的值域有同样的维数. 在以后, 当我们建立起 T 与 T^* 的矩阵表示之间的联系之后, 这将转为非常神秘的事实, 即一个 $m \times n$ 矩阵的行向量的线性生成的维数与其列向量的线性生成的维数是相同的, 这给了我们一个矩阵秩的概念. 在第五章, 我们将研究一种情况 (希尔伯特空间), 其中我们给出 V 与 V^* 之间一个固定的基本同构. 若 $T \in \mathrm{Hom}V$, 则当然 $T^* \in \mathrm{Hom}V^*$, 及我们可以利用这个同构 "移送"T^* 至 $\mathrm{Hom}V$. 但现在 T 可以与它的 (移送的) 伴随 T^* 相比较, 而它们可能相等. 即 T 可能自伴. 自伴变换是 "好" 的变换, 正如我们自己在简单情况中所看到的, 及很幸运地, 从理论物理引起的许多重要的线性映射都是自伴的.

若 $T \in \mathrm{Hom}(V, W)$ 及 $l \in W^*$, 则当然 $l \circ T \in V^*$. 而且, 由第一章, 定理 3.3 的推论可知映射 $l \mapsto l \circ T$(T 固定) 是一个由 W^* 至 V^* 的线性映射. 这个映射被称为 T 的伴随并被记为 T^*. 因此 $T^* \in \mathrm{Hom}(W^*, V^*)$ 及对于所有 $l \in W^*, T^*(l) = l \circ T$.

定理 3.4 *映射 $T \mapsto T^*$ 是由向量空间 $\mathrm{Hom}(V, W)$ 至向量空间 $\mathrm{Hom}(W^*, V^*)$ 的一个同构. 在定义域与值域的有关假定之下, 我们亦有 $(T \circ S)^* = S^* \circ T^*$.*

证明 通过 $T \mapsto T^*$ 的线性性, 我们在上面所说的每一件事都是 $\omega(l, T) = l \circ T$ 的双线性性的一个推论. 我们称之为 T^* 的映射仅为 ω_T, 所以由第一章定理 6.1 即得出 $T \mapsto T^*$ 的线性性. 读者可能再次受益于由 $(c_1T_1 + c_2T_2)^*(l)$ 开始的一个直接的线性性验证.

欲知 $T \mapsto T^*$ 是单射的, 我们取任意 $T \neq 0$ 并选取 $\alpha \in V$ 使 $T(\alpha) \neq 0$. 我们再选取 $l \in W^*$ 使 $l(T(\alpha)) \neq 0$. 由于 $l(T(\alpha)) = (T^*(l))(\alpha)$, 所以我验证了 $T^* \neq 0$.

其次, 若 $d(V) = m$ 及 $d(W) = n$, 则由定理 3.1 的推论亦有 $d(V^*) = m$ 及 $d(W^*) = n$, 又由定理 2.5 可知 $d(\mathrm{Hom}(V, W)) = mn = d(\mathrm{Hom}(W^*, V^*))$. 因此单射映射 $T \mapsto T^*$ 是一个同构 (由定理 2.4 的推论).

最后, $(T \circ S)^* l = l \circ (T \circ S) = (l \circ T) \circ S = S^*(l \circ T) = S^*(T^*(l)) = (S^* \circ T^*)l$, 所以 $(T \circ S)^* = S^* \circ T^*$. □

读者可能会猜想在 V 与 V^{**} 等同之下, T^{**} 变成与 T 等同起来. 这是的, 这是称同构 $\xi \mapsto \xi^{**}$ 为自然的真正理由. 我们将在本节末回到这个问题上来, 同时, 我们记录一个重要的基本恒等式.

定理 3.5 $(R(T^*))^\circ = N(T)$ 及 $N(T^*) = (R(T))^\circ$.

证明 当出现下面语句时, 它们定义地成对等价: $l \in N(T^*)$, $T^*(l) = 0, l \circ T = 0$, 对于所有 $\xi \in V$ 有 $l(T(\xi)) = 0$, $l \in (R(T))^\circ$. 因此 $N(T^*) = (R(T))^\circ$. 另一个的证明是类似的, 留给读者来做. [从 $\alpha \in N(T)$ 开始及结束于 $\alpha \in (R(T^*))^\circ$.] □

一个线性变换的*秩*为它的值空间的维数.

推论 3.3 *T^* 的秩等于 T 之秩.*

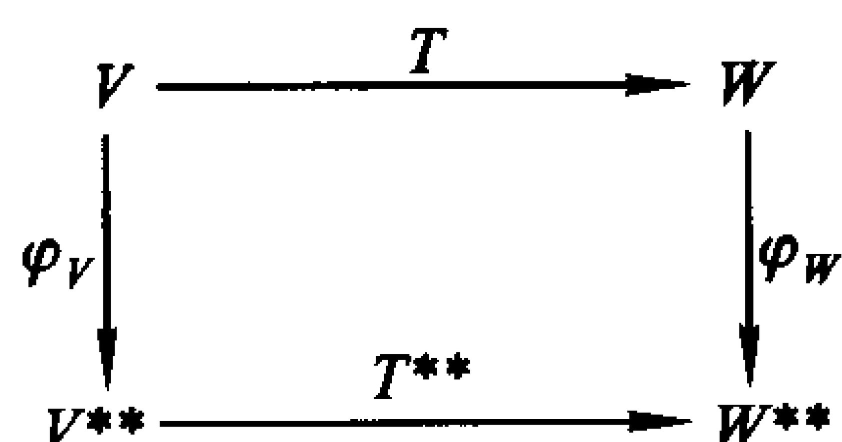
证明 由定理 2.4 与 3.3 可知 $R(T)$ 与 $(N(T))^\circ$ 之维数各为 $d(V) - d(N(T))$, 及由上面定理可知第二个等于 $d(R(T^*))$. 所以, $d(R(T)) = d(R(T^*))$. □

并向量 考虑 $\mathrm{Hom}(V, W)$ 中任意 T, 它的值域是一维的. 若 β 是 M 中一个非零向量, 则 $x \mapsto x\beta$ 是一个基同构 $\theta : \mathbb{R} \to M$ 及 $\theta^{-1} \circ T : V \to R$ 是一个线性泛函 $\lambda \in V^*$. 所以 $T = \theta \circ \lambda$ 及对于所有 ξ, $T(\xi) = \lambda(\xi)\beta$. 我们记这个为 $T = \lambda(\cdot)\beta$, 并称任意这种 T 为一个*并向量*.

引理 3.2 *若 T 为并向量 $\lambda(\cdot)\beta$, 则 T^* 是并向量 $\beta^{**}(\cdot)\lambda$.*

证明 $(T^*(l))(\xi) = (l \circ T)(\xi) = l(T(\xi)) = l(\lambda(\xi)\beta) = l(\beta)\lambda(\xi)$, 所以 $T^*(l) = l(\beta)\lambda = \beta^{**}(l)\lambda$, 及 $T^* = \beta^{**}(\cdot)\lambda$. □

* **再论自然同构** 我们现在可以更精细地来阐述自然同构的概念. 由上面我们看到由一个有限维向量空间 V 到它第二对偶的所有同构之中, 我们可以自然地挑出一个, 即映射 $\xi \mapsto \xi^{**}$, 此处对于所有 V^* 中 f 有 $\xi^{**}(f) = f(\xi)$. 让我们称这个同构为 φ_V.'自然' 这个字的学术含义在于所有这些同构的集合 $\{\varphi_V\}$, 我们发现了一种对于每一个空间 V 选取一个同构 φ_V 的方式, 并发现证明这是一个 "自然的" 选择是由于各种 φ_V 是互相关联着的这种顺当的方式. 欲看出关于这一点的含意, 我们考虑两个有限维空间 V 与 W 及一个 $\mathrm{Hom}(V, W)$ 中的映射 T. 则 $T^* \in \mathrm{Hom}(W^*, V^*)$ 与 $T^{**} = (T^*)^* \in \mathrm{Hom}(V^{**}, W^{**})$. 关于四个映射 T, T^{**}, φ_V 与 φ_W 的安置可以被显示一张图如下:



这个图指出了由 V 至 W^{**} 的两个映射 $\varphi_W \circ T$ 与 $T^{**} \circ \varphi_V$, 若对于任意 V, W 与 T, 这两个映射总是相等的, 则我们定义同构 $\{\varphi_V\}$ 的集合为自然的. 这是两种方法围绕图形行进得到同样结果的条件, 即图形是可交换的.

当 V 被等同于 V^{**}(由 φ_V) 及 W 被等同于 W^{**}(被 φ_W), 则这是 T"变成"T^{**} 的条件之另一途径. 我们将证明留作一个习题.

习题

3.1 命 θ 为由一个向量空间 V 至 $\mathbb{R}^n$ 的一个同构. 试证泛函 $\{\pi_i \circ \theta\}_1^n$ 构成 V^* 的一个基.

3.2 试证由 $\mathbb{R}^n$ 至 $(\mathbb{R}^n)^*$ 的标准同构, 即我们由 $\mathbb{R}^n$ 的标准基 (恒等元) 的坐标同构与 $(\mathbb{R}^n)^*$ 的对偶基同构之复合而得到的, 恰好是我们的朋友 $\boldsymbol{a} \mapsto l_{\boldsymbol{a}}$, 此处 $l_{\boldsymbol{a}}(\boldsymbol{x}) = \sum_1^n a_i x_i$(试证对偶基同构为 $\boldsymbol{a} \mapsto \sum_1^n a_i \pi_i$).

3.3 由定理 1.6 我们得知 V 的一个基 $\{\beta_i\}$ 的选取定义了一个从 W^n 至 $\mathrm{Hom}(V, W)$ 的同构, 这里 W 为任何向量空间. 应用这个事实及定理 1.3 得到 V^* 的一个基, 并试证这个基是 $\{\beta_i\}$ 的对偶基.

3.4 试证课文中列举的关于 A° 的性质.

3.5 试找出 $\mathbb{R}^3$ 中 $\prec 1,1,1 \succ$ 的零化子 (的一个基).(利用 $(\mathbb{R}^3)^*$ 与 $\mathbb{R}^3$ 的同构将基向量表为三元组.)

3.6 试找出 $\mathbb{R}^3$ 中 $\{\prec 1,1,1 \succ, \prec 1,2,3 \succ\}$ 的零化子 (的一个基).

3.7 试找出 $\mathbb{R}^4$ 中 $\{\prec 1,1,1,1 \succ, \prec 1,2,3,4 \succ\}$ 的零化子 (的一个基).

3.8 试证若 $V = M \bigoplus N$, 则 $V^* = M^\circ \bigoplus N^\circ$.

3.9 试证若 M 为一个 n 维向量空间 V 的任意子空间及 $d(M) = m$, 则 M 可以被看成 V 的 m 个元素的一个独立子集合的线性生成或 V^* 的 $n-m$ 个元素的一个独立子集合 (的零空间的交) 的零化子.

3.10 若 $B = \{f_i\}_1^m$ 为 $V(B \subset V^*)$ 上线性泛函的一个有限集合, 则它的零化子仅为泛函 f_i 的零空间 $N_i = N(f_i)$ 的交 $N = \bigcap_1^n N_i$. 试陈述本课文中定理 3.3 的对偶. 即取 W 为泛函 f_i 的线性生成, 所以 $W \subset V^*$ 及 $W^\circ \subset V$. 陈述推论的对偶.

3.11 试证下面定理是定理 3.3 的推论的一个结果.

定理 3.6 命 N 为 V 上线性泛函的一个集合 $\{f_i\}_1^n$ 的零空间的交 $\bigcap_1^n N_i$, 并假定 V^* 中的 g 在 N 上为零. 则 g 为集合 $\{f_i\}_1^n$ 的一个线性组合.

3.12 定理 3.3 的一个推论为：若 W 为 V 的一个真子空间, 则至少存在 V^* 中一个非零线性泛函 f 使在 W 上有 $f = 0$. 试用初等方法直接地证明这个事实 (你被允许去构造一个适当的基.)

3.13 一个向量空间 V 上的一个线性泛函的 m 元组 $\{f_i\}_1^m$ 定义了一个由 V 到 $\mathbb{R}^m$ 的线性映射 $\alpha \mapsto \prec f_1(\alpha), \cdots, f_m(\alpha) \succ$. 这里用到了什么定理? 试证这个线性映射的值域是整个 $\mathbb{R}^m$ 当且仅当 $\{f_i\}_1^m$ 是一个泛函的独立集合.[提示: 若值域是一个真子空间 W, 则存在一个非零 m 元组 $\boldsymbol{a}$ 使对于所有 $\boldsymbol{x} \in W, \sum_1^n a_i x_i = 0$.]

3.14 继续上面的习题, 试问什么是线性映射 $\alpha \mapsto \prec f_1(\alpha), \cdots, f_m(\alpha) \succ$ 的零空间 N? 若 g 为一个线性泛函, 它在 N 上取值零, 现在作为上面习题及第一章定理 4.3 的一个推论, 试证 g 是 f_i 的一个线性组合.(假定集合 $\{f_i\}_1^m$ 是独立的.)

3.15 详细写出 T^* 是线性的之证明 [对于给定 $\mathrm{Hom}(V, W)$ 中之 T]. 同时直接地证明 $T \mapsto T^*$ 是线性的.

3.16 试证定理 3.5 的另外一半.

3.17 当 $i = 1, 2$ 时, 命 θ_i 为由 V_i 至 V_i^{**} 的同构 $\alpha \mapsto \alpha^{**}$, 并假定给定 $T \in \mathrm{Hom}(V_1, V_2)$. 不严谨的语句 $T = T^{**}$ 的精确含义为

$$T^{**} = \theta_2 \circ T \circ \theta_1^{-1} \quad \text{或} \quad T^{**} \circ \theta_1 = \theta_2 \circ T.$$

试证这个恒等式, 像通常那样, 我们用证明它对于每一个 $\alpha \in V_1$ 皆成立来论证它.

3.18 命 $\theta : \mathbb{R}^n \to V$ 为一个基同构. 试证若 $(\mathbb{R}^n)^*$ 以自然的方式等同于 $\mathbb{R}^n$, 则伴随 θ^* 是对偶基的坐标同构.

3.19 命 ω 为 $V \times W$ 上的任意双线性泛函. 则两个相关的线性变换为由 $(T(\xi))(\eta) = \omega(\xi, \eta)$ 定义的 $T : V \to W^*$ 及由 $(S(\eta))(\xi) = \omega(\xi, \eta)$ 定义的 $S : W \to V^*$. 试证若 W 等同于 W^{**}, 则 $S = T^*$.

3.20 假定 $f \in (\mathbb{R}^m)^*$ 有坐标 m 元组 $\boldsymbol{a}[f(\boldsymbol{y}) = \sum_1^m a_i y_i]$ 及 $T \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 有矩阵 $\boldsymbol{t} = \{t_{ij}\}$. 试利用所有这些坐标写出数 $f(T(\boldsymbol{x}))$ 的显式表达式. 试重新安排和, 使它以形式

$$g(\boldsymbol{x}) = \sum_1^n b_i x_i$$

出现, 然后读出用 $\boldsymbol{a}$ 表示 $\boldsymbol{b}$ 的公式.

2.4 矩阵

矩阵与线性变换 在第一章中, 读者已经学到了关于矩阵与它跟线性变换关系的一些事情; 我们更系统的讨论将从复习这些早先的材料开始. 按照通常的概念, 矩阵是一个数的长方形阵, 如

$$\begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & & \vdots \\ t_{m1} & t_{m2} & \cdots & t_{mn} \end{bmatrix}.$$

注意第一个指标数是行指标而第二个指标数是列指标. 如果在这个阵列中共有 m 行及 n 列, 则称这是一个 $(m \times n)$ 矩阵. 这个概念是欠精确的. 一个长方形阵是画出一个矩阵的一种方式, 但一个矩阵真正地是一个函数, 恰如一个序列是一个函数. 利用记号 $\overline{m} = \{1, \cdots, m\}$, 则上面的矩阵就是一个函数, 对于 $\overline{m} \times \overline{n}$ 中的每一个整数对 $\prec i, j \succ$ 都给出一个数. 所以它是集合 $\mathbb{R}^{\overline{m} \times \overline{n}}$ 中的一个元素. 两个 $m \times m$ 矩阵的加法按照明显的位置到位置的方式来进行, 它仅仅是 $\mathbb{R}^{\overline{m} \times \overline{n}}$ 中两个函数的加法; 对于标量乘法同样成立. 所以所有 $m \times n$ 矩阵的集合是向量空间 $\mathbb{R}^{\overline{m} \times \overline{n}}$, 即具有一个有点想像的有限指标集合的笛卡儿空间. 对于 $\prec i, j \succ$ 上的函数 t 的值 $t(i, j)$, 我们将采用习惯的指标记号 t_{ij}, 及我们亦将 t 记为 $\{t_{ij}\}$, 这正如我们对序列及其他加标集合所做的一样.

矩阵的其他性质源于 $m \times n$ 矩阵与变换 $T \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 之间的对应性.

下面的定理重述了第一章的结果. 见第一章的定理 1.2, 1.3 与 6.2, 及 1.6 节末关于线性组合映射的讨论.

定理 4.1 命 $\{t_{ij}\}$ 为一个 $m \times n$ 矩阵, 及当 $j = 1, \cdots, n$ 时, 命 t^j 为 m 元组, 它为矩阵的第 j 列. 则存在一个惟一的 $T \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 使骨架 $T = \{t^j\}$, 满足 $T(\delta^j) = t^j$ 对所有 j 成立. T 被定义为线性组合映射 $x \mapsto y = \sum_{j=1}^n x_j t^j$, T 的一个等价表示为标量方程组

$$y_i = \sum_{j=1}^n t_{ij} x_j, \quad i = 1, \cdots, m.$$

$\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 中每一个 T 都从这个方式引起, 及由 $\mathbb{R}^{\overline{m} \times \overline{n}}$ 至 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 的双射. $\{t_{ij}\} \mapsto T$ 是一个自然同构.

这里需要作的仅有的注记为: 将一个 $m\times n$ 矩阵与一个 m 元组的 n 元组等同起来时, 我们使用了对偶性的标准等同性之一 (见 0.10 节). 我们处理着真正不同的空间 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 与 $(\mathbb{R}^{\overline{m}})^{\overline{n}}$ 之间的自然同构, 好像它们是恒等的.

我们也可以用 $\{t_{ij}\}$ 的行的方式使 T 与 $\{t_{ij}\}$ 相关联. 如上, 取 m 元组方程 $\boldsymbol{y}=\sum_{j=1}^{n}x_j\boldsymbol{t}^j$ 的第 i 个坐标, 则得等价而熟悉的数值 (标量的) 方程组 $y_i=\sum_{j=1}^{n}t_{ij}x_j(i=1,\cdots,m)$. 现在由 $\mathbb{R}^n$ 至 $\mathbb{R}$ 的映射 $\boldsymbol{x}\mapsto\sum_{j=1}^{n}c_jx_j$ 是 $\mathbb{R}^n$ 上最一般的线性泛函. 所以在上面的数值方程组中, 我们仅用到了矩阵 $\{t_{ij}\}$ 的 m 个行去给出 $\mathbb{R}^n$ 上线性泛函的 m 元组, 由第一章的定理 3.6 可知它等价于 $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$ 中单个 m 元组值的线性映射 T.

任意有限维空间 V 与 W 的有序基的选择允许我们将上面定理转至 $\mathrm{Hom}(V,W)$. 由于我们现在使 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 中一个矩阵 $\boldsymbol{t}$ 相关于 $\mathrm{Hom}(V,W)$ 中一个变换 T, 所以我们将上面讨论的 $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$ 中的变换记为 $\overline{T}$.

定理 4.2 命 $\{\alpha_j\}_1^n$ 与 $\{\beta_i\}_1^m$ 分别为向量空间 V 与 W 的有序基. 对于 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 中的每一个矩阵 $\{t_{ij}\}$, 命 T 为 $\mathrm{H}(V,M)$ 中惟一的元素, 它满足 $T(\alpha_j)=\sum_{i=1}^{m}t_{ij}\beta_i(j=1,\cdots,n)$. 则映射 $\{t_{ij}\}\mapsto T$ 是一个由 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 至 $\mathrm{Hom}(V,W)$ 的同构.

证明 我们仅仅联合上面定理中的同构 $\{t_{ij}\}\mapsto T$ 与由 $\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$ 至 $\mathrm{Hom}(V,W)$ 的同构 $\overline{T}\mapsto T=\psi\circ\overline{T}\circ\varphi^{-1}$, 此处 φ 与 ψ 为两个给定的基同构. 则 T 为定理中所描述的变换, 这是由于 $T(\alpha_j)=\psi(\overline{T}(\varphi^{-1}(\alpha_j)))=\psi(\overline{T}(\delta^j))=\psi(\boldsymbol{t}^j)=\sum_{i=1}^{m}t_{ij}\beta_i$. 映射 $\{t_{ij}\}\mapsto T$ 是两个同构之合成, 所以是一个同构. □

用略微不同的方式观察一下我们刚刚做过的事是有益的. 给定矩阵 $\{t_{ij}\}$, 命 τ_j 是 W 中的向量, 它的坐标 m 元组是矩阵的第 j 列 $\boldsymbol{t}^j$, 所以 $\boldsymbol{\tau}_j=\sum_{i=1}^{m}t_{ij}\beta_i$. 则命 T 为 $\mathrm{Hom}(V,W)$ 中惟一的元素, 它满足 $T(\alpha_j)=\boldsymbol{\tau}_j(j=1,\cdots,n)$. 现在我们由下面的两步从 $\{t_{ij}\}$ 得到 T: 在定理 1.6 给出的由 $\mathrm{Hom}(V,W)$ 至 W^n 的同构之下, T 对应于 n 元组 $\{\boldsymbol{\tau}_j\}_1^n$, 及由 W 与 $\mathbb{R}^m$ 之间的坐标同构延拓为它的由 W^n 至 $(\mathbb{R}^m)^n$ 之乘积同构而得 $\{\boldsymbol{\tau}_j\}_1^n$ 对应于矩阵 $\{t_{ij}\}$.

推论 4.1 若 $\boldsymbol{y}$ 为 W 中向量 η 的坐标 m 元组及 $\boldsymbol{x}$ 为 V 中 ξ 的坐标 n 元组 (关于给定基), 则 $\eta=T(\xi)$ 当且仅当 $y_i=\sum_{j=1}^{n}t_{ij}x_j(i=1,\cdots,m)$.

证明 我们知道标量方程等价于 $\boldsymbol{y}=\overline{T}(\boldsymbol{x})$, 它是方程 $\boldsymbol{y}=\psi^{-1}\circ T\circ\varphi(\boldsymbol{x})$. 同构 ψ 转换这个方程为方程 $\eta=T(\xi)$. □

我们现在的问题为去寻找线性变换之间的关系的矩阵类似. 对于笛卡儿空间 $\mathbb{R}^n$ 之间的变换, 这是相当直接而不困难的事, 这是因为我们知道这里矩阵对于变换是自然的自我替代. 但是, 当我们离开笛卡儿空间时, 一个变换 T 就不再由任何自然的方式得到矩阵, 而当基已被选好与笛卡儿空间的一个对应的 $\overline{T}$ 被得到时, 才得到一个矩阵. 现在所有矩阵都是关于选取基而被决定的, 所有的计算因必须出现基与坐标同构而都是复杂的. 有两种方式来处理这种情况. 第一种办法, 也是常用的方法为对一般空间 V 直接去描述事物, 并且就去接受涉及基与对偶基的、相应不太透明的、必然比较复杂的语句. 另一种可能性为首先读出关于笛卡儿空间的答案, 然后再经过坐标同构将它们重写出来.

引理 4.1 矩阵元素 t_{kj} 可以由 T 通过公式

$$t_{kj} = \mu_k(T(\alpha_j))$$

而得到, 此处 μ_k 为 W^* 中对偶基的第 k 个元素.

证明 $\mu_k(T(\alpha_j)) = \mu_k(\sum_{i=1}^m t_{ij}\beta_i) = \sum_i t_{ij}\mu_k(\beta_i) = \sum_i t_{ij}\delta_i^k = t_{kj}$. □

用笛卡儿空间的术语, $\overline{T}(\delta^j)$ 是 $\overline{T}$ 的矩阵 $\{t_{ij}\}$ 的第 j 个列 m 元组 t^j, 及 t_{kj} 为 t^j 的第 k 个坐标. 由线性映射的观点, 第 k 个坐标是应用第 k 个坐标投影 π_k 而得到的, 所以 $t_{kj} = \pi_k(\overline{T}(\delta^j))$. 在基同构之下, π_k 变成了 μ_k, $\overline{T}$ 变成了 T, δ^j 变成了 α_j, 及笛卡儿恒等变成了引理之恒等.

转置 $m \times n$ 矩阵 $\{t_{ij}\}$ 的转置是由 $t_{ij}^* = t_{ji}$(对所有 i, j) 定义的 $n \times m$ 矩阵 $\{t_{ij}^*\}$. t^* 的行当然就是 t 的列, 且其逆亦真.

定理 4.3 关于 W^* 与 V^* 中对偶基的 T^* 的矩阵就是 T 的矩阵的转置.

证明 若 s 为 T^* 的矩阵, 则由引理 3.1 与 4.1 可知

$$\begin{aligned} s_{ji} &= \alpha_{jj}^{**}(T^*(\mu_i)) = \alpha_j^{**}(\mu_i \circ T) \\ &= (\mu_i \circ T)(\alpha_j) = \mu_i(T(\alpha_j)) = t_{ij}. \end{aligned}$$ □

定义 4.1 矩阵 $\{t_{ij}\} \in \mathbb{R}^{\overline{m} \times \overline{n}}$ 的行空间为由 m 个行向量生成的 $\mathbb{R}^n$ 的子空间. 类似地, 列空间是由 $\mathbb{R}^m$ 中 n 个列向量生成的.

推论 4.2 一个矩阵的行与列空间有相同的维数.

证明 若 T 为由 $T(\delta^j) = t^j$ 定义的 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 的元素, 则矩阵 $\{t_{ij}\}$ 中列向量的集合 $\{t^j\}_1^n$ 是在 T 之下, $\mathbb{R}^n$ 的标准基之像, 所以是它的生成, 这就是我们所谓的矩阵的列空间, 它恰为 T 之值域. 特别列空间的维数等于 $d(R(T)) = \mathrm{rank}\, T$.

由于 T^* 的矩阵为矩阵 $\boldsymbol{t}$ 的转置 $\boldsymbol{t}^*$, 类似地, 我们得知 T^* 的秩为 $\boldsymbol{t}^*$ 的列空间的维数. 但 $\boldsymbol{t}^*$ 的列空间是 $\boldsymbol{t}$ 的行空间, 因此推论的结论归结为恒等式 $\mathrm{rank}\, T^* = \mathrm{rank}\, T$, 这就是定理 3.5 的推论. □

这个共同的维数被称为矩阵的*秩*.

矩阵乘积 若 $T \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 及 $S \in \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^l)$, 则当然 $R = S \circ T \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^l)$, 由分别表示 S 与 T 的矩阵 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 来计算 R 的矩阵 $\boldsymbol{r}$ 的确应该是可能的. 为了作这个计算, 我们置 $\boldsymbol{y} = T(\boldsymbol{x})$ 与 $\boldsymbol{z} = S(\boldsymbol{y})$, 所以 $\boldsymbol{z} = (S \circ T)(\boldsymbol{x}) = R(\boldsymbol{x})$. 用矩阵 $\boldsymbol{t}$ 与 $\boldsymbol{s}$ 表示的等价标量方程为

$$y_i = \sum_{h=1}^{n} t_{ih} x_h \quad 与 \quad z_k = \sum_{i=1}^{m} s_{ki} y_i,$$

所以

$$z_k = \sum_{i=1}^{m} s_{ki} \sum_{h=1}^{n} t_{ih} x_h = \sum_{h=1}^{n} \left(\sum_{i=1}^{m} s_{ki} t_{ih} \right) x_h.$$

但是 $z_k = \sum_{h=1}^{n} r_{kh} x_h (k = 1, \cdots, l)$. 取 $\boldsymbol{x}$ 为 δ^j, 我们得

$$r_{kj} = \sum_{i=1}^{m} s_{ki} t_{ij}, \quad 对于所有 k 与 j 成立.$$

因此我们找到了映射 $R = S \circ T : \boldsymbol{x} \to \boldsymbol{z}$ 的矩阵 $\boldsymbol{r}$ 的公式. 当然, $\boldsymbol{r}$ 被定义为矩阵 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 的乘积, 我们记为 $\boldsymbol{r} = \boldsymbol{s} \cdot \boldsymbol{t}$ 或 $\boldsymbol{r} = \boldsymbol{st}$.

注意为了使乘积 $\boldsymbol{st}$ 可以被定义, 左边因子的列的个数必须等于右边因子的行的个数. 我们由 $\boldsymbol{s}$ 的第 k 行及同时 $\boldsymbol{t}$ 的第 j 列, 如我们所进行的, 将它们的对应元素乘起来, 再将乘积的结果加起来即得元素 r_{kj}. 这个过程表述于图 2.1 之中. 用 $\mathbb{R}^n$ 上标量积 $(\boldsymbol{x}, \boldsymbol{y}) = \sum_1^n x_i y_i$ 的语言, 我们看到 $\boldsymbol{r} = \boldsymbol{st}$ 中的元素 r_{kj} 是 $\boldsymbol{s}$ 的第 k 行与 $\boldsymbol{t}$ 的第 j 列的标量积.

由于我们已经定义了两个矩阵的乘积为它们对应变换的乘积之矩阵, 即映射 $T \mapsto \{t_{ij}\}$ 保持乘积 $(S \circ T \mapsto \boldsymbol{st})$, 所以由第一章的定理 4.1 的一般原则可知变换的合成适合的代数定律将自动地对矩阵的乘积成立. 例如, 我们未作一个具体的计算就知道矩阵的乘法是可结合的. 对于方阵, 我们有下面的定理.

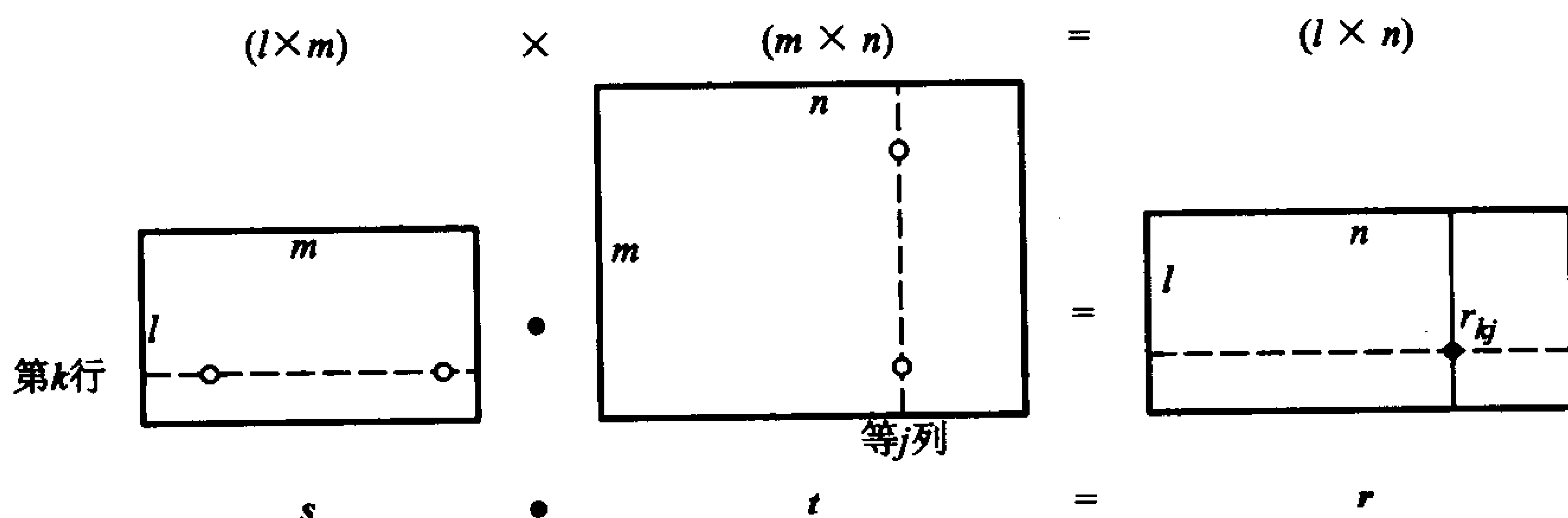


图 2.1

定理 4.4 $n \times n$ 方阵的集合 M_n 是一个代数，它自然地同构于代数 $\mathrm{Hom}(\mathbb{R}^n)$.

证明 我们已知 $T \mapsto \{t_{ij}\}$ 是由 $\mathrm{Hom}(\mathbb{R}^n)$ 至 M_n 的一个自然线性同构 (定理 4.1), 且我们已经定义了矩阵的乘积, 所以映射亦保持乘法. 所以我们由关于在 $\mathrm{Hom}(\mathbb{R}^n)$ 中成立的第一章的定理 3.5 的观察得知代数 (对于一个代数) 的定律对 M_n 成立. □

$\mathrm{Hom}(\mathbb{R}^n)$ 中的恒等元 I 取基向量 δ^j 至自身, 所以它的矩阵 $\boldsymbol{e}$ 有 δ^j 为它的第 j 列: $\boldsymbol{e}^j = \delta^j$. 因此当 $i = j$ 时有 $e_{ij} = \delta_i^j = 1$; 当 $i \neq j$ 时, $e_{ij} = \delta_i^j = 0$. 即矩阵 $\boldsymbol{e}$ 在其主对角线上有值 1(从左上角至右下角), 而其他位置皆为 0. 由于在代数同构 $T \mapsto \boldsymbol{t}$ 之下, $I \mapsto \boldsymbol{e}$, 所以我们知道在矩阵乘法中, $\boldsymbol{e}$ 是单位元. 当然我们可以直接地来验证: $\sum_{j=1}^{n} t_{ij}e_{jk} = t_{ik}$, 类似地将 $\boldsymbol{e}$ 乘在左边. 记号 '$\boldsymbol{e}$' 是含糊的, 我们曾用它表示对于任意 n, $n \times n$ 方阵空间 $\mathbb{R}^{\overline{n} \times \overline{n}}$ 中的单位.

推论 4.3 一个 $n \times n$ 方阵 $\boldsymbol{t}$ 有一个乘法逆元当且仅当它的秩为 n.

证明 由定理可知存在一个 $\boldsymbol{s} \in M_n$ 使 $\boldsymbol{st} = \boldsymbol{ts} = \boldsymbol{e}$ 当且仅当存在一个 $S \in \mathrm{Hom}(\mathbb{R}^n)$ 使 $S \circ T = T \circ S = I$. 但这样一个 S 存在当且仅当 T 是一个同构, 及由定理 2.4 的推论可知这等价于 T 的值域的维数为 n. 但这个维数等于 $\boldsymbol{t}$ 的秩, 证完. □

如果一个方阵 (或 $\mathrm{Hom}V$ 中一个变换) 是可逆的, 则称它为非奇异的.

定理 4.5 若 $\{\alpha_i\}_1^n$, $\{\beta_j\}_1^m$, 及 $\{\gamma_k\}_1^l$ 分别为向量空间 U, V, 与 W 的有序基, 又若 $T \in \mathrm{Hom}(U, V)$ 与 $S \in \mathrm{Hom}(V, W)$, 则 $S \circ T$ 的矩阵为 S 与 T(关于给定基)

的矩阵的乘积.

证明 由定义可知 $S \circ T$ 的矩阵为 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^l)$ 中, $\overline{S \circ T} = \chi^{-1} \circ (S \circ T) \circ \varphi$ 的矩阵, 此处 φ 与 χ 为 U 与 W 的给定基同构. 但若 ψ 为 V 的基同构, 则

$$\overline{S \circ T} = (\chi^{-1} \circ S \circ \psi) \circ (\psi^{-1} \circ T \circ \varphi) = \overline{S} \circ \overline{T},$$

所以由矩阵乘法的定义可知它的矩阵是 $\overline{S}$ 与 $\overline{T}$ 的矩阵之乘积. 后者是 S 与 T 关于给定基的矩阵, 将上述观察综合起来即得定理. □

矩阵乘积与转置之间有一个简单的关系.

定理 4.6 *若矩阵乘积 $\boldsymbol{st}$ 已定义, 则 $\boldsymbol{t}^*\boldsymbol{s}^*$ 亦然, 于是 $\boldsymbol{t}^*\boldsymbol{s}^* = (\boldsymbol{st})^*$.*

证明 一个直接的计算是容易的. 我们有

$$(st)^*_{jk} = (st)_{kj} = \sum_{i=1}^{m} s_{ki}t_{ij} = \sum_{i=1}^{m} t^*_{ji}s^*_{ik} = (t^*s^*)_{jk}.$$

因此如所宣称的 $(\boldsymbol{st})^* = \boldsymbol{t}^*\boldsymbol{s}^*$. □

这个恒等式显然是变换恒等式 $(S \circ T)^* = T^* \circ S^*$ 的矩阵形式, 及如果需要, 它可以由后面恒等式推导出来.

笛卡儿向量作为矩阵 我们可以将 n 元组 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ$ 或者看成一个 $n \times 1$ 矩阵, 在这种情况下, 我们称它为一个*列向量*; 或者可替代地看成一个 $1 \times n$ 矩阵, 在这种情况下, 我们称它为一个*行向量*. 当然这种恒等性是自然同构. 这样做的部分原因为方程 $y_i = \sum_{j=1}^n t_{ij}x_j$ 确切地说为列向量 $\boldsymbol{y}$ 为 $\boldsymbol{t}$ 与列向量 $\boldsymbol{x}$ 的矩阵积, 即 $\boldsymbol{y} = \boldsymbol{tx}$. *当 $\mathbb{R}^n$ 被看成 $n \times 1$ 列向量空间时, 线性变换 $T: \mathbb{R}^n \to \mathbb{R}^m$ 变成了左乘以一个固定的矩阵 $\boldsymbol{t}$.* 基于这个原因, 我们将取列向量为一个 n 元组 $\boldsymbol{x}$ 的标准矩阵表示; 则 $\boldsymbol{x}^*$ 为对应的行向量.

特别, 一个线性泛函 $F \in (\mathbb{R}^n)^*$ 变成了它的矩阵的左边乘法, 这当然是 $1 \times n$(F 为由 $\mathbb{R}^n$ 至 $\mathbb{R}^1$), 所以这仅仅是 $\mathbb{R}^n$ 中一个 n 元组的行矩阵表示. 即在由 $\mathbb{R}^n$ 至 $(\mathbb{R}^n)^*$ 的自然同构 $\boldsymbol{a} \mapsto L_{\boldsymbol{a}}$ 中, 此处 $L_{\boldsymbol{a}}(\boldsymbol{x}) = \sum_1^n a_ix_i$, $L_{\boldsymbol{a}}$ 可以被解释为将一个 n 元组 $\boldsymbol{a}$ 看成行向量 $\boldsymbol{a}^*$ 的左矩阵乘法. 行向量 ($1 \times n$ 矩阵)$\boldsymbol{a}^*$ 与列向量 ($n \times 1$ 矩阵)$\boldsymbol{x}$ 的矩阵乘积为一个 1×1 矩阵 $\boldsymbol{a}^* \cdot \boldsymbol{x}$, 即一个数.

现在让我们看看关于 T^*, 这些观察能得到什么. 数 $L_{\boldsymbol{a}}(T(\boldsymbol{x}))$ 是 1×1 矩阵 $\boldsymbol{a}^*\boldsymbol{tx}$. 由 T^* 的定义可知 $L_{\boldsymbol{a}}(T(\boldsymbol{x})) = (T^*(L_{\boldsymbol{a}}))(\boldsymbol{x})$, 所以我们得知泛函 $T^*(L_{\boldsymbol{a}})$

是行向量 a^*t 的左乘积. 由于 L_a 的行向量形式为 a^*, $T^*(L_a)$ 的行向量形式为 a^*t, 所以这证明了, 当 $\mathbb{R}^n$ 上的泛函被解释为行向量时, T^* 就变成了 t 的右乘积. 这仅仅是对我们已知的某些事的重复. 我们如果取转置使行向量转成 n 元组的标准列向量, 则它证明了 T^* 为 t^* 的左乘积, 所以这给出了 T^* 的矩阵为 t^* 的另一证明.

基的变换 若 $\varphi: x \mapsto \xi = \sum_1^n x_i\beta_i$ 与 $\theta: y \mapsto \xi = \sum_1^n y_i\beta_i'$ 为 V 的两个基同构, 则 $A = \theta^{-1} \circ \varphi$ 为 $\mathrm{Hom}(\mathbb{R}^n)$ 中的同构, 它将一个向量 ξ 关于基 $\{\beta_i\}$ 的坐标 n 元组 x 变为同样向量关于基 $\{\beta_i'\}$ 的坐标 n 元组 y. 同构 A 被称为 "坐标变换" 同构. 利用 A 的矩阵 a, 我们有 $y = ax$ 如上.

坐标变换映射 $A = \theta^{-1}\varphi$ 不应该与看来类似的 $T = \theta \circ \varphi^{-1}$ 相混淆. 后者是 V 上的一个映射, 它是 $\mathrm{Hom}(V)$ 的一个元素, 将每一个 β_i 映至 β_i'.

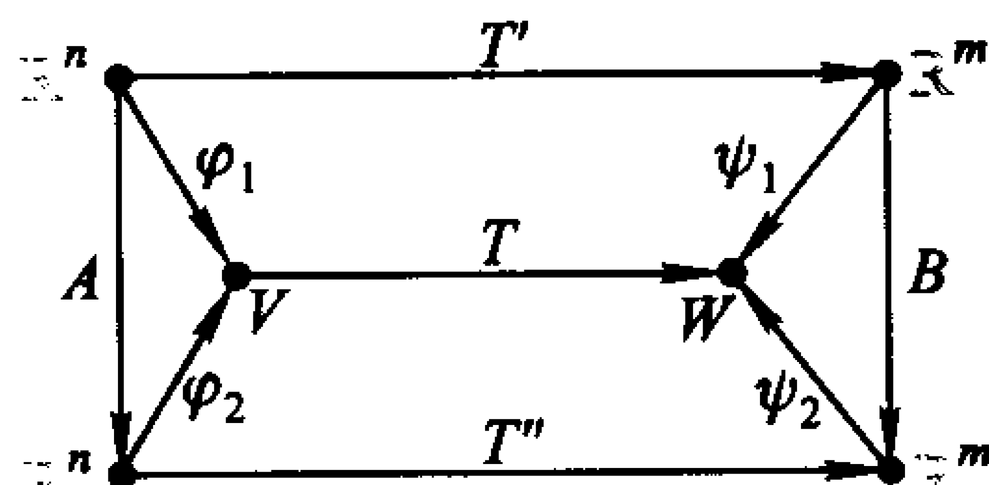


图 2.2

我们现在要看看当我们在其定义域与值域空间变换基时, 一个变换 $T \in \mathrm{Hom}(V, W)$ 的矩阵会发生什么. 假定 φ_1 与 φ_2 为由 $\mathbb{R}^n$ 至 V 的基同构, 及 ψ_1 与 ψ_2 为由 $\mathbb{R}^m$ 至 W 的基同构, 及 t' 与 t'' 分别为 T 关于第一个与第二个基的矩阵. 即 t' 为 $T' = (\psi_1)^{-1} \circ T \circ \varphi_1 \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 的矩阵, 类似地有 t'' 之含义. 映射 $A = \varphi_2^{-1} \circ \varphi_1 \in \mathrm{Hom}(\mathbb{R}^n)$ 是关于 V 的坐标变换映射: 若 x 是一个向量 ξ 关于第一个基的坐标 n 元组 [即 $\xi = \varphi_1(x)$], 则 $A(x)$ 为关于第二个基的坐标 n 元组. 类似地, 命 B 为 W 的坐标变换映射 $\psi_2^{-1} \circ \psi_1$. 则图 2.2 将有助于保持这些空间与映射的各种直线关系. 我们称这个图是可交换的, 它是指两个点之间的两条道路表示同样的映射, 选择不同的路线对, 我们就可以读出 9 个映射 $T, T', T'', \varphi_1, \varphi_2, A_1, \psi_1, \psi_2, B$ 之间的所有恒等式. 例如, T'' 可以先沿 A 之逆向走, 然后沿 T', 最后沿 B 走而得到. 即 $T'' = B \circ T' \circ A^{-1}$. 因为这些 "外圈映射" 都是笛卡儿空间的映射, 所以我们可以读出对应的矩阵恒等式

$$t'' = bt'a^{-1},$$

这显示了 T 关于第二对基的矩阵如何从它关于第一对基的矩阵而得到.

在读出上面恒等式时, 我们真正做的事是从图中的长路径里由定义消去某些折回的步骤. 因此由定义, 我们得到

$$B\circ T'\circ A^{-1}=(\psi_2^{-1}\circ\psi_1)\circ(\psi_1^{-1}\circ T\circ\varphi_1)\circ(\varphi_1^{-1}\circ\varphi_2)=\psi_2^{-1}\circ T\circ\varphi_2=T''.$$

在上面的情况中, 定义域与上域空间是不同的, 而两个基变换是彼此独立的. 若 $W=V$, 则 $T\in\mathrm{Hom}(V)$, 当然我们只要考虑一下基变换. 于是, 公式变成

$$\boldsymbol{t}''=\boldsymbol{a}\cdot\boldsymbol{t}'\cdot\boldsymbol{a}^{-1}.$$

现在我们考虑一个线性泛函 $F\in V^*$. 若 $\boldsymbol{f}''$ 与 $\boldsymbol{f}'$ 为被考虑为列向量($n\times 1$ 矩阵) 它的坐标 n 元组, 则 F 关于这两个基的矩阵为行向量 $(\boldsymbol{f}')^*$ 与 $(\boldsymbol{f}'')^*$, 恰如我们较早见到的. 在值空间中没有基的变化, 这是因在此 $W=\mathbb{R}$, 且它的永久自然基向量为 1. 所以, 在公式 $\boldsymbol{t}''=\boldsymbol{b}\boldsymbol{t}'\boldsymbol{a}^{-1}$ 中 $\boldsymbol{b}=\boldsymbol{e}$, 因而我们有 $(\boldsymbol{f}'')^*=(\boldsymbol{f}')^*\boldsymbol{a}^{-1}$ 或

$$\boldsymbol{f}''=(\boldsymbol{a}^{-1})^*\boldsymbol{f}'.$$

我们需要将这个结果与一个向量 $\xi\in V$ 的坐标变换相比较, 正如我们较早所见, 它由

$$\boldsymbol{x}''=\boldsymbol{a}\boldsymbol{x}'$$

给出. 这些变换按反向进行 (投入转置). 大部分出于历史原因, V^* 中的泛函 F 被称为协变向量,由于 V 中一个坐标的变换为 V^* 中对应坐标变换的矩阵之逆的转置, 所以 V 中向量 ξ 被称为反变向量. 这些名词在经典张量分析与微分几何中被用到.

由笛卡儿空间 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 对应的同构 $\{t_{ij}\}\mapsto T$ 自动地为一个基同构. 它在 $\mathrm{Hom}(V,W)$ 中的基是在 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 中标准基的同构之下的像, 此处后者是指克罗内克函数 δ^{kl} 的集合, 它的定义为: 当 $\prec k,l\succ\neq\prec i,j\succ$ 时, $\delta^{kl}(i,j)=0$, 及 $\delta^{kl}(k,l)=1$.(回忆在 $\mathbb{R}^A$ 中, δ^a 为满足当 $b\neq a$ 时, $\delta^a(b)=0$, 及 $\delta^a(a)=1$ 之函数. 在此 $A=\overline{m}\times\overline{n}$, 而 A 中元素 a 则为有序对 $a=\prec k,l\succ$.) 函数 δ^{kl} 为矩阵, 除第 l 个列之外, 其他列皆为 0, 而第 l 列为 m 元组 δ^k. 对应的变换 D_{kl} 除 α_l 之外, 它将每一个基向量 α_j 变至 0, 而将 α_l 变至 β_k. 即当 $j\neq l$ 时, $D_{kl}(\alpha_j)=0$, $D_{kl}(\alpha_l)=\beta_k$. 同样, D_{kl} 将 V 中第 l 个基向量变为 W 的第 k 个基向量, 而将 V 中其他基向量变至 0.

若 $\xi=\sum x_i\alpha_i$, 则 $D_{kl}(\xi)=x_l\beta_k$.

由于 $\{D_{kl}\}$ 是由同构 $\{t_{ij}\} \mapsto T$ 定义的基, 所以 $\{t_{ij}\}$ 是 T 关于这个基的坐标集合; 它是在坐标同构之下, T 的像. 很有兴趣地去看看 T 关于这个基的展开是怎样自动地出现的. 我们有

$$T(\xi) = T\left(\sum_{j=1}^{n} x_j\alpha_j\right) = \sum_{j=1}^{n} x_j T(\alpha_j) = \sum_{i,j} t_{ij}x_j\beta_i = \sum_{i,j} t_{ij}D_{ij}(\xi),$$

所以

$$T = \sum_{i,j} t_{ij}D_{ij}.$$

我们关于 V^* 中对偶基的原始讨论是现在情况的特例. 我们有 $\mathrm{Hom}(V,\mathbb{R}) = V^*$, 它有 $\mathbb{R}$ 的永久标准基 1. V^* 的基对应于 V 的基 $\{\alpha_i\}$, 所以含有那些映射 D_l, 它将 α_l 变为 1, 而当 $j \neq l$ 时, 则将 α_j 变为 0. 因此 $D_l(\xi) = D_l(\sum x_j\alpha_j) = x_l$, 及 D_l 为第 l 个坐标泛函 ε_l.

最后, 我们注意到 $T \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 的矩阵表示对于我们较早在 1.5 节中讨论过 T 的区组分解是非常有建议性的. 在习题中, 我们将要求读者去证明事实上有 $T_{kl} = t_{kl}D_{kl}$.

习题

4.1 试证若 $\omega : V \times V \to \mathbb{R}$ 为 V 上一个双线性泛函, 及 $T : V \to V^*$ 为由 $(T(\eta))(\xi) = \omega(\xi,\eta)$ 定义的对应线性变换, 则对于 V 的任意基 $\{\alpha_i\}$, 矩阵 $t_{ij} = \omega(\alpha_i, \alpha_j)$ 为 T 之矩阵.

4.2 试验证下面矩阵的行与列的秩均为 1.

$$\begin{bmatrix} -5 & 2 & 3 \\ -10 & 4 & 6 \end{bmatrix}.$$

4.3 试用直接计算证明若一个 2×3 矩阵的行秩为 1, 则其列秩亦为 1.

4.4 命 $\{f_i\}_1^3$ 为 $\mathbb{R}$ 上 $\mathfrak{C}^2$ 函数 (两次连续可微实值函数) 的一个线性互依集合. 试证对于任何 x, 三个 3 元组 $\prec f_i(x), f_i'(x), f_i'(x) \succ$ 是互依的. 所以证明 $\sin t, \cos t$ 与 e^t 是线性独立的 (计算一个选好 x 的微商三元组).

4.5 试计算

$$\begin{bmatrix} 5 & -2 & 3 & 1 \\ -4 & 1 & 2 & -3 \\ 1 & 6 & -1 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ -3 & 0 \\ 4 & 2 \end{bmatrix}.$$

4.6 试计算

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

并根据你的答案给出

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$$

存在的充要条件.

4.7 如果 $\boldsymbol{a}^2 = \boldsymbol{a}$, 则称矩阵 $\boldsymbol{a}$ 为*幂等的*. 试找出所有 2×2 矩阵的包含所有幂等矩阵的向量空间 $\mathbb{R}^{2\times 2}$ 之一个基.

4.8 试由直接计算证明

$$\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$$

是可逆的并找出其逆.

4.9 试用显式地解方程

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\cdot\begin{bmatrix} x & y \\ z & w \end{bmatrix}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

来证明左端之矩阵为可逆的当且仅当 (行列式)$ad-bc$ 不为零.

4.10 试找一个 2×2 非零矩阵

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

它的平方为零.

4.11 试找出所有 2×2 矩阵, 它们的平方为零.

4.12 试用计算矩阵乘积证明矩阵乘法是可结合的.

4.13 类似地, 试直接地证明分配律 $(\boldsymbol{r}+\boldsymbol{s})\cdot\boldsymbol{t}=\boldsymbol{r}\cdot\boldsymbol{t}+\boldsymbol{s}\cdot\boldsymbol{t}$.

4.14 试证用一个 $\mathbb{R}^{\overline{m}\times\overline{n}}$ 中的固定矩阵 $\boldsymbol{r}$ 之左边乘积是一个由 $\mathbb{R}^{\overline{n}\times\overline{p}}$ 至 $\mathbb{R}^{\overline{m}\times\overline{p}}$ 的线性变换. 第一章中什么定理做对应的这件事?

4.15 试证两个矩阵乘积的秩最多为它们秩的最小者.(回忆一个矩阵的秩为它对应的 T 的值空间之维数.)

4.16 命 $\boldsymbol{a}$ 是一个 $m\times n$ 矩阵及 $\boldsymbol{b}$ 为一个 $n\times m$ 矩阵. 若 $m>n$, 试证 $\boldsymbol{a}\cdot\boldsymbol{b}$ 不可能为单位 $\boldsymbol{e}(m\times m)$.

4.17 命 Z 为形如

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

的 2×2 矩阵子集合. 试证 Z 是一个 $\mathbb{R}^{2\times 2}$ 的子代数(即 Z 在加法、标量乘法与矩阵乘法之下是闭的). 再证事实上, Z 同构于复数系.

4.18 若一个矩阵 (必须是方阵) 等于它的转置, 则称为对称的. 作为方形阵列, 它关于主对角线是对称的. 试证对于任意 $m\times n$ 矩阵 $\boldsymbol{t}$, 乘积 $\boldsymbol{t}\cdot\boldsymbol{t}^*$ 有意义且为对称的.

4.19 试证若 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 为 $n\times n$ 对称矩阵又若它们是可交换的, 则 $\boldsymbol{s}\cdot\boldsymbol{t}$ 是对称的 (不要试图用写出矩阵乘积来回答这个问题). 再证逆命题, 即若 $\boldsymbol{s},\boldsymbol{t}$ 与 $\boldsymbol{s}\cdot\boldsymbol{t}$ 都是对称的, 则 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 可交换.

4.20 假定 $T\in \mathrm{Hom}\mathbb{R}^2$ 有一个对称矩阵及 T 不是形如 cI 的矩阵, 试证 T 正好有两个特征向量 (至一个标量倍数). 关于 $\mathbb{R}^2$ 包含这两个特征向量的 "特征基", T 的矩阵变成什么?

4.21 试证 2×2 对称矩阵 $\boldsymbol{t}$ 有一个对称平方根 $\boldsymbol{s}(\boldsymbol{s}^2=\boldsymbol{t})$ 当且仅当它的特征值是非负的 (假定上面习题).

4.22 假定 $\boldsymbol{t}$ 为一个 2×2 矩阵, 满足 $\boldsymbol{t}^*=\boldsymbol{t}^{-1}$. 试证 $\boldsymbol{t}$ 为下面形式之一

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix},\quad \begin{bmatrix} a & b \\ b & -a \end{bmatrix},$$

此处 $a^2+b^2=1$.

4.23 试证乘以上面的 $\boldsymbol{t}$ 是欧氏等距的. 即证明若 $\boldsymbol{y}=\boldsymbol{t}\cdot\boldsymbol{x}$, 此处 $\boldsymbol{x}$ 与 $\boldsymbol{y}\in\mathbb{R}^2$, 则 $\|\boldsymbol{x}\|=\|\boldsymbol{y}\|$, 其中 $\|\boldsymbol{x}\|=(x_1^2+x_2^2)^{1/2}$.

4.24 命 $\{D_{kl}\}$ 为课文中定义的 $\mathrm{Hom}(V,W)$ 的基. 取 $W=V$, 试证这些算子满足非常重要的乘法规则

$$D_{ij}\circ D_{kl}=0,\quad \text{当}j\neq k,$$
$$D_{ik}\circ D_{kl}=D_{il}.$$

4.25 将上面的恒等式记在心上, 试证若 $l\neq m$, 则存在 $\mathrm{Hom}V$ 中的变换 S 与 T 满足

$$S\circ T-T\circ S=D_{lm}.$$

同时, 寻找 S 与 T 满足

$$S\circ T-T\circ S=D_{ll}-D_{mm}.$$

4.26 给定 $T\in\mathrm{Hom}\mathbb{R}^n$, 由第一章我们知道 $T=\sum_{i,j}T_{ij}$, 此处 $T_{ij}=P_iTP_j$ 及 $P_i=\theta_i\pi_i$. 现在我们亦有

$$T=\sum_{k,l}t_{kl}D_{kl}$$

试由课文中 D_{ij} 的定义证明 $P_iD_{ij}P_j=D_{ij}$ 及当 $i\neq k$ 或 $j\neq l$ 时有 $P_iD_{kl}P_j=0$. 因此 $T_{ij}=t_{ij}D_{ij}$.

2.5 迹与行列式

这一个短节的目的为让读者通晓 Hom$V$ 上两个非常特别的实值函数, 并陈述它们的一些性质.

定理 5.1 若 V 是一个 n 维向量空间, 则在向量空间 $\mathrm{Hom}(V)$ 上恰好存在一个线性泛函 λ 使对于所有 $\mathrm{Hom}(V)$ 中的 S,T 皆具有 $\lambda(S\circ T)=\lambda(T\circ S)$ 的性质, 且正则化使 $\lambda(I)=n$. 若 V 的一个基被选取及 T 的对应矩阵为 $\{t_{ij}\}$, 则 $\lambda(T)=\sum_{i=1}^n t_{ii}$, 即主对角线上元素之和.

证明 如果我们选择一个基及定义 $\lambda(T)$ 为 $\sum_1^n t_{ii}$, 则显然 λ 是 $\mathrm{Hom}(V)$ 上的一个线性泛函及 $\lambda(I)=n$. 而且,

$$\lambda(S\circ T)=\sum_{i=1}^n\left(\sum_{j=1}^n s_{ij}t_{ji}\right)=\sum_{i,j=1}^n s_{ij}t_{ji}=\sum_{i,j}t_{ji}s_{ij}=\lambda(T\circ S).$$

即 V 的每一个基皆给我们一个 $(\mathrm{Hom}V)^*$ 中的泛函 λ 满足 $\lambda(S\circ T)=\lambda(T\circ S),\lambda(I)=n$, 及 $\lambda(T)=\sum t_{ii}$ 对于该基的矩阵表示成立.

现在假定 μ 为 $(\mathrm{Hom}(V))^*$ 中任意元素满足 $\mu(S\circ T)=\mu(T\circ S)$ 与 $\mu(I)=n$. 如果我们选取 V 的一个基并利用由 $\mathbb{R}^{\overline{n}\times\overline{n}}$ 至 Hom$V$ 的同构 $\theta:\{t_{ij}\}\mapsto T$, 则我们得到 $\mathbb{R}^{\overline{n}\times\overline{n}}$ 上一个泛函 $\nu=\mu\circ\theta(\nu=\theta^*\mu)$ 满足 $\nu(\boldsymbol{st})=\nu(\boldsymbol{ts})$ 及 $\nu(\boldsymbol{e})=n$. 由定理 4.1(或 3.1) 可知 ν 由一个矩阵 $\boldsymbol{c}$ 给出, $\nu(\boldsymbol{t})=\sum_{i,j=1}^n c_{ij}t_{ij}$, 及方程 $\nu(\boldsymbol{st}-\boldsymbol{ts})=0$ 变成了 $\sum_{i,j,k=1}^n c_{ij}(s_{ik}t_{kj}-s_{jk}t_{ki})=0$.

我们将它留给读者作为一个习题, 即证明当 $l\neq m$ 时, 则非常简单的特殊矩阵 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 可以被选出使这个和约化为 $c_{lm}=0$, 及用另一种选择, 约化为 $c_{ll}-c_{mm}=0$.

与要求 $\nu(\boldsymbol{e})=n$ 在一起, 这将导出当 $l\neq m$ 时, $c_{lm}=0$ 而当 $m=1,\cdots,n$ 时, $c_{mm}=1$. 即 $\nu(\boldsymbol{t})=\sum_1^n t_{mm}$, 及 ν 为用到的基 λ. 综合起来, 这证明了 $(\mathrm{Hom}V)^*$ 中存在惟一的 λ; 对于所有的 S 与 T, 满足 $\lambda(S\circ T)=\lambda(T\circ S)$ 与 $\lambda(I)=n$, 及在每一个基之下, $\lambda(T)$ 有如 $\sum t_{ii}$ 之对角线赋值. □

这个惟一的 λ 被称为迹泛函, 及 $\lambda(T)$ 为 T 的迹.它通常被记为 $\mathrm{tr}(T)$.

Hom$V$ 上的行列式函数 $\Delta(T)$ 就要复杂多了, 在第七章之前, 我们将不证明它的存在性. 它的几何含义如下: 首先, $|\Delta(T)|$ 是 T 乘以体积之因子. 更精确地说, 如果我们定义一个 V 的子集合的 "体积"ν, 它由一个选取的基及利用坐标

对应性将 $\mathbb{R}^n$ 上的"自然"体积移送至 V, 则对于任意图形 $A \subset V$, $\nu(T[A]) = |\Delta(T)| \cdot \nu(A)$. 这将在第八章中讲清楚. 第二, 按照 T 保持定向与否来确定 $\Delta(T)$ 为正的或负的, 这又是一个人为的概念, 我们将在以后来解释. 暂时, 我们将列举 $\Delta(T)$ 与这个几何表述相关的性质, 及给出一个表示 Δ 的惟一性的充分的数.

我们假定对于每一个有限维的向量空间 V, 皆存在一个由 $\mathrm{Hom}(V)$ 至 $\mathbb{R}$ 的函数 Δ(或当定义域有任何问题时, 则为 Δ_V) 使下面成立:

(a) 在 $\mathrm{Hom}(V)$ 中对任意 S,T, $\Delta(S \circ T) = \Delta(S)\Delta(T)$ 成立.

(b) 若 V 的一个子空间 N 在 T 之下是不变的及 T 在 N 与 V/N 上是恒等的 (即 $T[\overline{\alpha}] = \overline{\alpha}$ 对于 N 的每一个陪集 $\overline{\alpha} = \alpha + N$ 成立), 则 $\Delta(T) = 1$. 这样一个 T 是 V 的沿平行于 N 的平面的一个剪切. 二维的情况, 它可以被画出, 如图 2.3 所示.

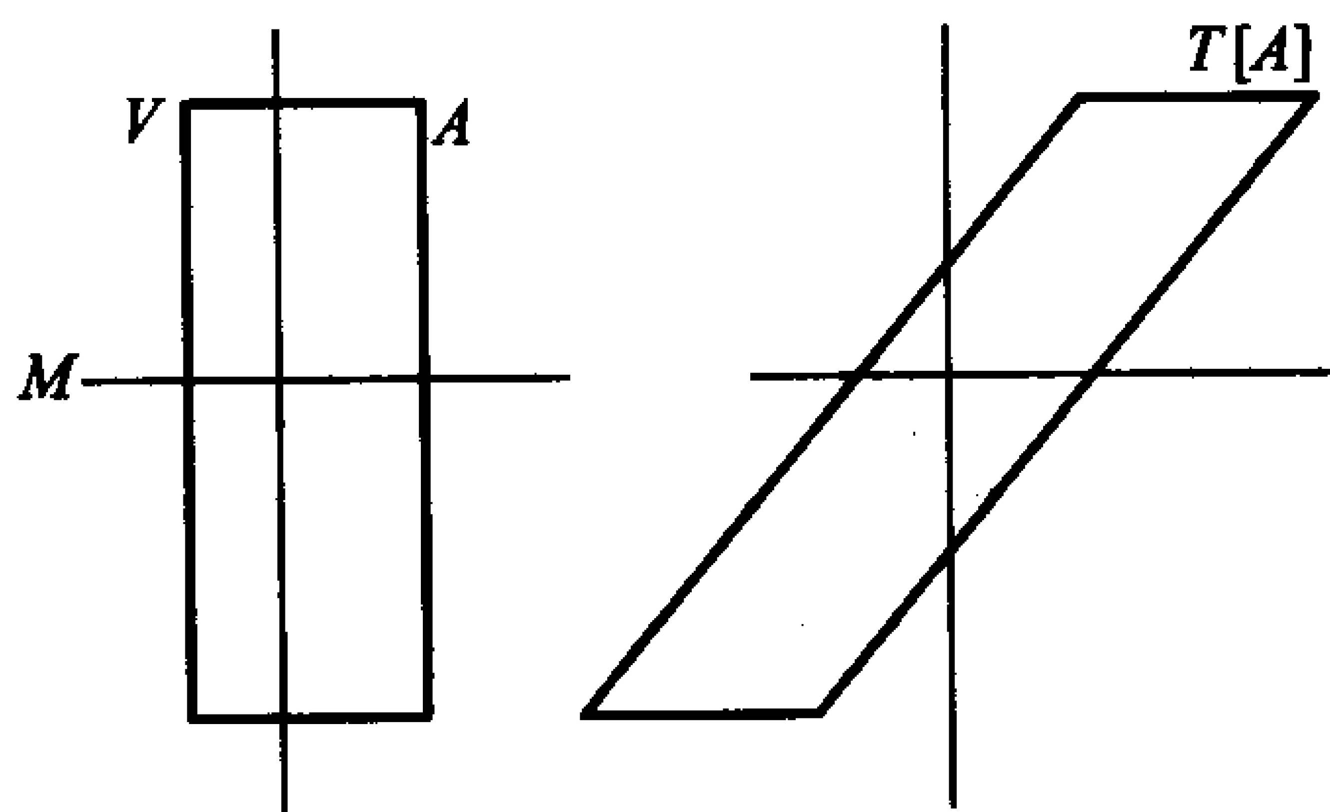


图 2.3

(c) 若 V 为 $T-$ 不变子空间 M 与 N 的一个直和 $V = M + N$, 又若 $R = T \restriction M$ 与 $S = T \restriction N$, 则 $\Delta(T) = \Delta(R)\Delta(S)$. 更准确地说, $\Delta_V(T) = \Delta_M(R)\Delta_N(S)$.

(d) 若 V 为一维的, 则任意 $T \in \mathrm{Hom}(V)$ 仅为一个常数 c_T 之乘积, 所以 $\Delta(T)$ 即为该常数.

(e) 若 V 为二维的及 T 交换一对独立向量, 则 $\Delta(T) = -1$. 这显然是一个单纯的定向变化性质.

Δ 由这些性质惟一地被确定这个事实将在下一节我们的讨论里得到, 它亦将给出计算 Δ 的一个程序. 这个程序对于维数大于 2 是有效的, 但对于 $T \in \mathrm{Hom}(\mathbb{R}^2)$, $\Delta(T)$ 有一个简单的公式, 每一个学生都应记在心中.

定理 5.2 若 $T \in \mathrm{Hom}(\mathbb{R}^2)$ 及 $\{t_{ij}\}$ 是它的 2×2 矩阵, 则 $\Delta(T) = t_{11}t_{22} - t_{12}t_{21}$.

这是一般公式的一个特例, 我们将在第七章中推导出 $\Delta(T)$ 表达为 $n!$ 项的和, 其中每一项都是 T 的矩阵的 n 个数的一个积. 这个公式对于大的 n, 在计算上太复杂了而不用, 但对于 $n=3$, 它与下一节中用我们的行约化计算差不多容易, 而对于 $n=2$, 它就变成了上面的简单表示式. 还有 Δ 的一些性质, 每一个学生都应该熟悉的. 它们皆将在第七章中来证明.

定理 5.3 若 $T \in \mathrm{Hom}V$, 则 $\Delta(T^*) = \Delta(T)$. 若 θ 为由 V 至 W 的一个同构及 $S = \theta \circ T \circ \theta^{-1}$, 则 $\Delta(S) = \Delta(T)$.

定理 5.4 变换 T 是非奇异的 (可逆的) 当且仅当 $\Delta(T) \neq 0$.

在下一个定理中, 我们考虑 $T \in \mathrm{Hom}\mathbb{R}^n$, 及我们要设想 $\Delta(T)$ 为 T 的矩阵 $\boldsymbol{t}$ 的一个函数. 为了强调这一点, 我们使用记号 $D(\boldsymbol{t}) = \Delta(T)$.

定理 5.5 (克拉默法则) 给定一个 $n \times n$ 矩阵 $\boldsymbol{t}$ 及一个 n 元组 $\boldsymbol{y}$, 命 $\boldsymbol{t}|_j\boldsymbol{y}$ 为将 $\boldsymbol{t}$ 的第 j 列换成 $\boldsymbol{y}$ 所得到的矩阵. 则对于所有 j, 我们有

$$\boldsymbol{y} = \boldsymbol{t} \cdot \boldsymbol{x} \Rightarrow D(\boldsymbol{t})x_j = D(\boldsymbol{t}|_j\boldsymbol{y}).$$

若 $\boldsymbol{t}$ 为非奇异的 $[D(\boldsymbol{t}) \neq 0]$, 则这就变成了方程 $\boldsymbol{y} = \boldsymbol{t} \cdot \boldsymbol{x}$ 的解 $\boldsymbol{x}$ 的显式公式; 尽管在一些情况下, 它是不实用的 (大的 n), 但它在理论上却很重要.

习题

5.1 利用习题 4.25, 完成定理 5.1 的证明.

5.2 由我们关于迹的讨论可知 $\mathrm{tr}(T) = \sum t_{ii}$ 与基无关. 试证这个事实可以直接由

$$\mathrm{tr}(\boldsymbol{t} \cdot \boldsymbol{s}) = \mathrm{tr}(\boldsymbol{s} \cdot \boldsymbol{t})$$

及上一节的基变换公式中得出.

5.3 试用直接计算证明函数 $d(\boldsymbol{t}) = t_{11}t_{22} - t_{12}t_{21}$ 满足 $d(\boldsymbol{s} \cdot \boldsymbol{t}) = d(\boldsymbol{s})d(\boldsymbol{t})$ (此处 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 为 2×2 矩阵). 由此可知若 V 为二维的及对于 $T \in \mathrm{Hom}V$, $d(T)$ 由选择的一个基及 $d(T) = d(\boldsymbol{t})$ 来定义, 则 $d(T)$ 实际上是独立于基的.

5.4 继续上面习题, 试证对于下面任何情况皆有 $d(T) = \Delta(T)$:

(1) T 交换两个独立向量.

(2) T 有两个特征向量.

(3) T 有一个形如下面的矩阵

$$\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}.$$

再证若 T 不是上面形式, 则 $T = R \circ S$, 此处 S 为型 (1) 及 R 为型 (2) 或 (3).[提示: 假定 $T(\alpha) = \beta$, 其中 α 与 β 为独立的. 命 S 交换 α 与 β, 及考虑 $R = T \circ S$]. 最后证明对于所有 $T \in \mathrm{Hom}V$(V 是二维的), 我们有 $d(T) = \Delta(T)$.

5.5 若 $\boldsymbol{t}$ 为一个对称 2×2 矩阵, 试证存在一个 2×2 矩阵 $\boldsymbol{s}$ 满足 $\boldsymbol{s}^* = \boldsymbol{s}^{-1}, \Delta(\boldsymbol{s}) = 1$ 从而 $\boldsymbol{sts}^{-1}$ 为对角线矩阵.

5.6 假定定理 5.2, 试对于 2×2 之情况验证定理 5.4.

5.7 假定定理 5.2, 试对于 2×2 之情况验证定理 5.5.

5.8 在这个习题中, 我们假定读者知道一个实变量的连续函数是什么. 假定 2×2 矩阵函数

$$\boldsymbol{a}(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix}$$

有连续分量 $a_{ij}(t)(t \in (0,1))$, 并假定对于每一个 $t, \boldsymbol{a}(t)$ 都是非奇异的. 试证若 $x_1(t)$ 与 $x_2(t)$ 是连续的, 则线性方程 $\boldsymbol{a}(t)\cdot \boldsymbol{y}(t) = \boldsymbol{x}(t)$ 的解 $\boldsymbol{y}(t)$ 有连续的分量 $y_1(t)$ 与 $y_2(t)$.

5.9 一个齐次二阶线性微分方程是一个如下形式的方程

$$y'' + a_1 y' + a_0 y = 0.$$

此处 $a_1 = a_1(t)$ 与 $a_0 = a_0(t)$ 为连续函数. 一个解为一个 $\mathfrak{C}^2$ 函数 f(即一个二次连续可微函数), 它满足 $f''(t) + a_1(t)f'(t) + a_0(t)f(t) = 0$. 假定 f 与 g 为 $\mathfrak{C}^2$ 函数 (在 (0,1) 上), 满足 2×2 矩阵

$$\begin{bmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{bmatrix}$$

总是非奇异的. 试证存在一个齐次二阶微分方程使它们都是解.

5.10 在上面习题中, 试证所有解的空间是一个二维向量空间. 即证若 $h(t)$ 是任何第三个解, 则 h 是 f 与 g 的一个线性组合.

5.11 所谓笛卡儿平面 $\mathbb{R}^n$ 至自身的一个 "线性运动", 我们的意思是由 [0,1] 至 2×2 非奇异矩阵满足 $\boldsymbol{t}(0) = e$ 之集合的一个连续映射 $x \mapsto \boldsymbol{t}(x)$. 试证 $\Delta(\boldsymbol{t}(1)) > 0$.

5.12 试证若 $\Delta(\boldsymbol{s}) = 1$, 则存在一个线性运动, 其最终矩阵 $\boldsymbol{t}(1)$ 为 $\boldsymbol{s}$.

2.6 矩阵计算

读者在中学代数中学过的求解线性方程组的计算方法无疑是“逐次迭代消元法”. 第一个方程解出第一个未知量, 然后将第一个未知量的表达式代入其余方程, 因此在剩下的方程中消去了第一个未知量. 其次, 由第二个方程解出第二个未知量, 然后在剩下的方程中消去这个未知量. 按这个方法, 一次消去一个未知量, 最后得到一个解.

同样的方法亦解决了下面另外一些问题:

(1) 求得 $\mathbb{R}^n$ 中一个 m 个向量的集合的线性生成的一个显式基; 所以, 特别,

(2) 寻求这样一个子空间的维数;

(3) 计算一个 $m \times m$ 矩阵的行列式;

(4) 计算一个可逆 $m \times m$ 逆阵的逆.

在这一节中, 我们将简短地学习这个方法并解答这些问题.

我们首先注意到那些加于一个有限向量序列, 但不影响其生成的变换类型.

引理 6.1 命 $\{\alpha_i\}_1^m$ 为一个向量空间的任意向量的 m 元组, 又命 $\{\beta_i\}_1^m$ 为由 $\{\alpha_i\}_1^m$ 经任意下面初等运算之一而得到的:

(1) 交换两个向量;

(2) 用一个非零标量乘某一 α_i;

(3) 对于某些 $j \neq i$ 及某一 $x \in \mathbb{R}$, 将 α_i 换成 $\alpha_i - x\alpha_j$. 则

$$L(\{\beta_i\}_1^m) = L(\{\alpha_i\}_1^m).$$

证明 若 $\alpha_i' = \alpha_i - x\alpha_j$, 则 $\alpha_i = \alpha_i' + x\alpha_j$. 因此若 $\{\beta_i\}_1^m$ 是由型 (3) 的一个运算从 $\{\alpha_i\}_1^m$ 而得到的, 则 $\{\alpha_i\}_1^m$ 可以从 $\{\beta_i\}_1^m$ 由型 (3) 的一个运算而得到. 特别, 每一个序列都含于另一个的线性生成之中, 从而两个线性生成是相同的.

类似地, 每一个其他运算也可以由一个同类型的运算被还原, 所以线性生成是不变的. □

当我们在一个矩阵的行向量序列上应用这些运算时, 我们就称它们为*初等行运算*.

我们定义一个 n 元组 $x = \prec x_1, \cdots, x_n \succ$ 的级为第一个非零元素之指标. 因此, 若当 $i < j$ 时, $x_i = 0$, 而 $x_j \neq 0$, 则 x 的级为 j. $\prec 0,0,0,2,-1,0 \succ$ 的级为 4.

命 $\{a_{ij}\}$ 为一个 $m \times n$ 矩阵, 命 V 为它的行空间, 又命 $n_1 < n_2 < \cdots < n_k$ 为整数, 它们是 V 中非零向量之级. 今往构造 V 的一个基, 它包含 k 个元素, 正

好具有上面级的集合.

若 $\{a_{ij}\}$ 中每一个非零行皆有级 $> p$, 则 V 中每一个非零向量 $\boldsymbol{x}$ 有级 $> p$, 这是由于 $\boldsymbol{x}$ 是这些行向量的一个线性组合. 由于 V 中某向量有极小级 n_1, 所以 $\{a_{ij}\}$ 有某行, 它有级 n_1. 我们由行的变换将这样一行移至顶层. 然后我们将这个行 $\boldsymbol{x}$ 乘以一个常数, 使其第一个非零元素 x_{n_1} 为 1. 命 $\boldsymbol{a}^1, \cdots, \boldsymbol{a}^n$ 为我们现在有的行向量, 所以 $\boldsymbol{a}^1$ 有级 n_1 及 $a^1_{n_1} = 1$. 其次, 我们从其他行减去 $\boldsymbol{a}^1$ 的一个倍数使新的第 i 行的 n_1 坐标为 0. 特别地, 我们将 $\boldsymbol{a}^i$ 换成 $\boldsymbol{a}^i - a^i_{n_1} \cdot \boldsymbol{a}^1 (i > 1)$. 我们这样得到的矩阵有如下性质: 当 $j < n_1$ 时, 它的第 j 列为零 m 元组, 而它的第 n_1 列则为 $\mathbb{R}^m$ 中的 δ^1. 它的第一行有级 n_1, 而其他每一行都有级 $> n_1$. 它的行空间仍为 V, 我们仍称它为 $\boldsymbol{a}$.

现在命 $\boldsymbol{x} = \sum_1^m c_i \boldsymbol{a}^i$ 为 V 中一个级为 n_2 的向量. 则 $c_1 = 0$, 否则若 $c_1 \neq 0$, 则 $\boldsymbol{x}$ 的级为 n_1, 矛盾. 因此, $\boldsymbol{x}$ 是第 2 至第 m 行的线性组合. 如同第一种情况, 所以这些行中之一必定有级 n_2.

我们将上面的方法全部重复一次, 锁住这个向量. 我们将它换至第二行, 并使其 n_2 坐标为 1, 再从其他行减去它的某倍数 (包括第一行), 所以最终的矩阵的第 n_2 列为 δ^2. 其次, 我们找一个级为 n_3 的行, 并将它换到第三行, 再将第 n_3 列变成 δ^3, 等等.

我们由下面的 3×4 矩阵来展示这个方法. 这个例子的选取是随便的, 当使用 (2) 时, 分数并未出现. 当读者手算时, 将不会那样幸运. 我们的办法为保持矩阵简单并使方法本身易见.

$$\begin{bmatrix} 0 & -1 & 2 & 3 \\ 2 & 2 & 4 & -2 \\ 2 & 4 & 0 & 3 \end{bmatrix} \xrightarrow[(1)]{} \begin{bmatrix} 2 & 2 & 4 & -2 \\ 0 & -1 & 2 & 3 \\ 2 & 4 & 0 & 3 \end{bmatrix} \xrightarrow[(2)]{} \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & -1 & 2 & 3 \\ 2 & 4 & 0 & 3 \end{bmatrix}$$

$$\xrightarrow[(3)]{} \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & -1 & 2 & 3 \\ 0 & 2 & -4 & 5 \end{bmatrix} \xrightarrow[(2)]{} \begin{bmatrix} 1 & 1 & 2 & -1 \\ 0 & 1 & -2 & -3 \\ 0 & 2 & -4 & 5 \end{bmatrix}$$

$$\xrightarrow[(3)]{} \begin{bmatrix} 1 & 0 & 4 & 2 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 11 \end{bmatrix} \xrightarrow[(1)]{} \begin{bmatrix} 1 & 0 & 4 & 2 \\ 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{(3)} \begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

注意由最后的矩阵, 我们可以说出行空间的级为 1,2 与 4, 而原来的矩阵仅展示了级 1 与 2.

我们终止于一个 $m \times n$ 矩阵, 它有同样的行空间 V 及下面的特殊结构:

(1) 对于 $1 \leqslant j \leqslant k$, 第 j 行有级 n_j.

(2) 若 $k < m$, 则剩下的 $m-k$ 行为零 (因为一个非零行将有级 $> n_k$, 故得一个矛盾).

(3) 第 n_j 列为 δ^j.

因此, 前 k 行的任意线性组合, 其中系数为 $c_1, \cdots, c_k$, 则第 n_j 个位置为 c_j, 所以除非所有 c_j 皆为零外, 它是不为零的. 这 k 个向量形成 V 的一个基, 解决了问题 (1) 与 (2).

最后的矩阵被称为行约化梯形式. 可以证明它是被空间 V 与上面关于其行至 V 的元素之级的要求而惟一地被确定的. 它是 V 的标准基的行形式. 一个典型的行约化梯矩阵由图 2.4 所示. 这个矩阵是 8×11, 它的级为 1,4,5,7,10, 及其行空间有维数 5. 在折线之下全是 0. 前 5 行中的击点表示任意数, 但剩余元素的任意改变都会改变生成空间 V.

图 2.4

现在我们将从一般线性理论的观点来看看行约化运算的重要性. 在这个讨论中, 应用第 4 节中的事实将是方便的, 即若 $\mathbb{R}^n$ 中一个 n 元组被看成一个 $n \times 1$ 矩阵 (即一个列向量), 则线性方程组 $y_i = \sum_{j=1}^{n} a_{ij}x_j$, $i = 1, \cdots, m$ 正好表示单

个矩阵方程 $\boldsymbol{y} = \boldsymbol{a} \cdot \boldsymbol{x}$. 因此相关的线性变换 $A \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 现在就被看作仅仅被一个矩阵 $\boldsymbol{a}$ 相乘; $\boldsymbol{y} = A(\boldsymbol{x})$ 当且仅当 $\boldsymbol{y} = \boldsymbol{a} \cdot \boldsymbol{x}$.

我们首先注意, 加于一个 $m \times n$ 矩阵 $\boldsymbol{a}$ 上的每一个初等行运算皆等价于左乘一个对应的 $m \times m$ 初等矩阵$\boldsymbol{u}$. 暂时假定是这样, 则可以利用 $m \times m$ 单位矩阵 $\boldsymbol{e}$ 来找到 $\boldsymbol{u}$ 是什么. 因为 $\boldsymbol{u} \cdot \boldsymbol{a} = (\boldsymbol{u} \cdot \boldsymbol{e}) \cdot \boldsymbol{a}$, 所以我们得知将运算作用于矩阵 $\boldsymbol{a}$ 的结果可以由矩阵 $\boldsymbol{u} \cdot \boldsymbol{e}$ 左乘 $\boldsymbol{a}$ 而得到. 即若初等运算可以由 $\boldsymbol{u}$ 的矩阵乘法而得到, 则乘子为 $\boldsymbol{u} \cdot \boldsymbol{e}$. 这个论证建议我们应该将运算施于 $\boldsymbol{e}$, 然后看看若将所得的矩阵左乘以 $\boldsymbol{a}$, 所作的 $\boldsymbol{a}$ 上的运算.

若初等运算为交换第 i_0 与第 j_0 行, 则作用于 $\boldsymbol{e}$ 得到矩阵 $\boldsymbol{a}$, 其中当 $k \neq i_0$. 及 $k \neq j_0$ 时有 $u_{kk} = 1, u_{i_0 j_0} = u_{j_0 i_0} = 1$, 及对于所有其他指标有 $u_{kl} = 0$. 而且, 检查定义矩阵乘积 $\boldsymbol{u} \cdot \boldsymbol{a}$ 的元素之和将证明左乘这个 $\boldsymbol{u}$ 正好交换了任意 $m \times n$ 矩阵 $\boldsymbol{a}$ 的第 i_0 与第 j_0 行.

按同样方式可知, 将 $\boldsymbol{a}$ 的第 i_0 行乘以 c 等价于去左乘矩阵 $\boldsymbol{u}$, 除 $u_{i_0 i_0} = c$ 之外, 它与 $\boldsymbol{e}$ 一样. 最后, 第 j_0 行乘以 x 再加到第 i_0 行等价于左乘一个矩阵 $\boldsymbol{u}$, 除 $u_{i_0 j_0}$ 用 x 代替 0 之外, 皆与单位矩阵 $\boldsymbol{e}$ 一样.

这三个初等矩阵用图解法示于图 2.5 中. 每一个除指示的值外, 在主对角线上有值 1, 而在其余位置有值 0.

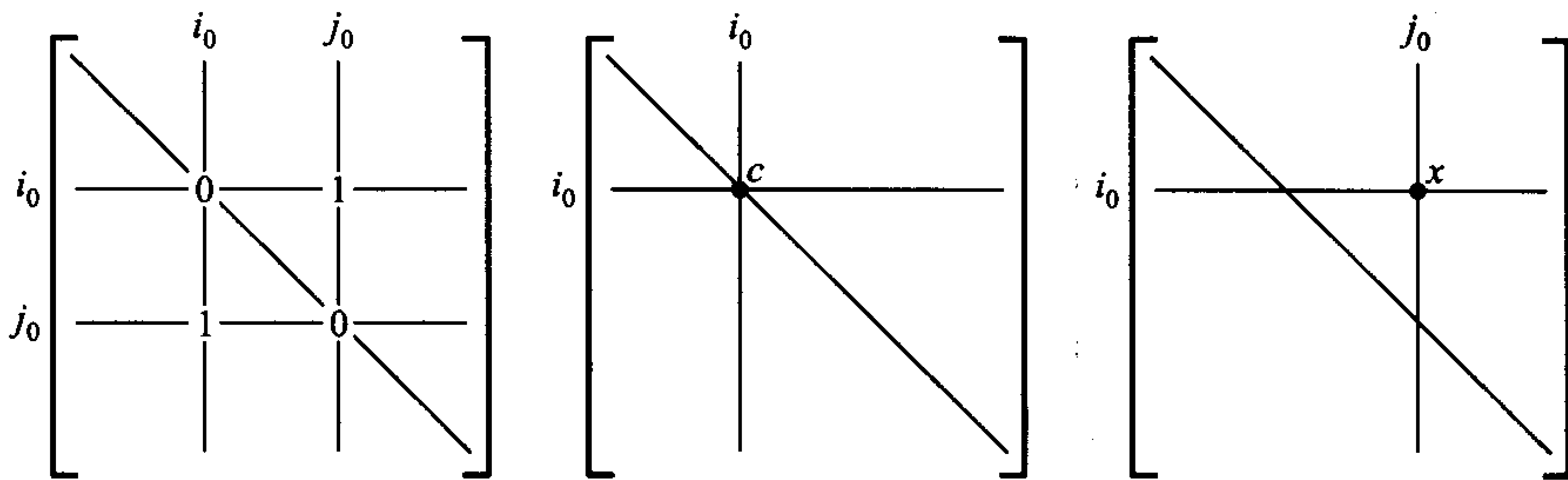


图 2.5

这些初等矩阵 $\boldsymbol{u}$ 都是非奇异的 (可逆的). 行交换矩阵是它自己的逆. 第 j 行乘以 c 之逆为同一行乘以 $1/c$. 结果将第 j 行的 c 倍加于第 i 行之逆为 $-c$ 乘以第 j 行加至第 i 行.

若 $\boldsymbol{u}^1, \boldsymbol{u}^2, \cdots, \boldsymbol{u}^p$ 为一个初等矩阵序列, 又若

$$\boldsymbol{b} = \boldsymbol{u}^p \cdot \boldsymbol{u}^{p-1} \cdot \cdots \cdot \boldsymbol{u}^1,$$

则 $\boldsymbol{b}\cdot\boldsymbol{a}$ 为由向 $\boldsymbol{a}$ 施以对应的在 $\boldsymbol{a}$ 上的初等行运算序列而得到的矩阵. 若 $\boldsymbol{u}',\cdots,\boldsymbol{u}^p$ 为行约化 $\boldsymbol{a}$ 的一个序列, 则 $\boldsymbol{r}=\boldsymbol{b}\cdot\boldsymbol{a}$ 为最终的行约化梯矩阵.

现在假定 $\boldsymbol{a}$ 是一个 $m\times m$ 方阵而且是非奇异的 (可逆的). 因此, 行空间的维数是 m, 所以有 m 个不同的级 $n_1,\cdots,n_k$. 即 $k=m$, 又由于 $1\leqslant n_1<\cdots<n_m=m$, 所以我们亦有 $n_i=i, i=1,\cdots,m$. 回忆 $\boldsymbol{r}$ 的第 n_i 列为 δ^i, 所以我们得知 $\boldsymbol{r}$ 的第 i 列为 δ^i, 因此 $\boldsymbol{r}$ 仅仅为单位矩阵 $\boldsymbol{e}$, 从而 $\boldsymbol{b}\cdot\boldsymbol{a}=\boldsymbol{e}$ 及 $\boldsymbol{b}$ 为 $\boldsymbol{a}$ 之逆.

让我们用这个程序找出

$$\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$$

之逆. 行约化序列为

$$\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\xrightarrow[(3)]{}\begin{bmatrix}1 & 2\\ 0 & -2\end{bmatrix}\xrightarrow[(2)]{}\begin{bmatrix}1 & 2\\ 0 & 1\end{bmatrix}\xrightarrow[(3)]{}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}.$$

对应的初等矩阵为

$$\begin{bmatrix}1 & 0\\ -3 & 1\end{bmatrix},\quad\begin{bmatrix}1 & 0\\ 0 & -1/2\end{bmatrix},\quad\begin{bmatrix}1 & -2\\ 0 & 1\end{bmatrix}.$$

因此逆为乘积

$$\begin{bmatrix}1 & -2\\ 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1/2\end{bmatrix}\begin{bmatrix}1 & 0\\ -3 & 1\end{bmatrix}=\begin{bmatrix}1 & 1\\ 0 & -1/2\end{bmatrix}\begin{bmatrix}1 & 0\\ -3 & 1\end{bmatrix}$$
$$=\begin{bmatrix}-2 & 1\\ 3/2 & -1/2\end{bmatrix}.$$

如果你有疑惑, 可以验证它.

最后, 由于 $\boldsymbol{b}\cdot\boldsymbol{e}=\boldsymbol{b}$, 我们看到我们由 $\boldsymbol{e}$, 应用同样的行运算, 即我们用过的约化 $\boldsymbol{a}$ 为梯形式 (集合在一起作为 $\boldsymbol{b}$ 来左乘) 而得到 $\boldsymbol{b}$. 这可能是计算一个矩阵的逆的最佳途径. 为了保持运算的路径, 我们可以将 $\boldsymbol{e}$ 置于 $\boldsymbol{a}$ 之右端而构成一个单个的 $m\times 2m$ 矩阵 $\boldsymbol{a}|\boldsymbol{e}$, 然后来行约化它. 在梯形式中, 它将为 $m\times 2m$ 矩阵 $\boldsymbol{e}|\boldsymbol{b}$, 所以我们就可以读出原来矩阵 $\boldsymbol{a}$ 之逆 $\boldsymbol{b}$.

让我们用这个方法重新计算一下

$$\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}$$

之逆. 我们行约化

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]$$

得到

$$\left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right] \xrightarrow[(3)]{} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{array}\right] \xrightarrow[(2)]{} \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right]$$

$$\xrightarrow[(3)]{} \left[\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & -1/2 \end{array}\right],$$

由此我们读出逆就是

$$\left[\begin{array}{cc} -2 & 1 \\ 3/2 & -1/2 \end{array}\right].$$

最后, 我们考虑计算一个 $m \times m$ 方阵的行列式的问题. 我们使用两个初等运算 (一个被修改过) 如下:

(1′) 交换两行, 并同时变更其中之一的符号;

(3) 如上, 对于某 $j \neq i$, 将某行 α_i 换为 $\alpha_i - x\alpha_j$.

当它们被用到一个方阵之行时, 这些运算使行列式不变. 这可以列举于第 5 节中行列式的性质得到, 其证明将留作一个习题. 而且, 这些性质将为我们在第七章中关于一个行列式的定义的平凡推论.

然后考虑一个 $m \times m$ 方阵 $\{a_{ij}\}$. 我们交换第一行与第 p 行使具有最小级 n_1 的行换至最上面的位置, 同时将向下移动的行变号 (在此就是第一行). 我们不要让新的第一行的首项系数为 1, 现在不用这个初等运算. 我们做减法, 从剩余行中减去第一行的某倍数, 使第 n_1 列的所有剩余元素皆为 0. 现在第 n_1 列就是 $c_1\delta^1$, 此处 c_1 为第一行的首项系数. 而新矩阵与原来的矩阵有同样的行列式.

按照上面的修改步骤, 我们继续上面的方法. 在一个变换中, 我们改变向下移动的一行的符号, 我们不去让 首项系数为 1, 及清除干净第 n_j 列使之变成 $c_j\delta^{n_j}$, 此处 c_j 为第 j 行的首项系数 ($1 \leqslant j \leqslant k$). 如前, 剩下的 $m-k$ 行为 0(若 $k < m$). 让我们称最终的矩阵为半约化的. 注意我们可以由此应用 k 次运算 (2) 以求得对应的约化梯矩阵; 只要当 $j = 1, \cdots, k$ 时, 在第 j 行上乘以 $1/c_j$. 若 s 是我们由 a 利用 (1′) 与 (3) 得到的半约化矩阵, 则我们将在下面证明它的行列式, 从而 a 的行列式也是一样,为主对角线上元素之乘积: $\prod_{i=1}^{m} s_{ii}$. 扼要地, 我们

可以用下法计算一个方阵 a 的行列式, 即利用运算 $(1')$ 与 (3) 将 a 化成一个半约化矩阵 s, 然后取 s 的主对角线上数的乘积.

若我们将这个方法用于

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

则得

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \xrightarrow[(3)]{} \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix} \xrightarrow[(3)]{} \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix},$$

于是行列式为 $1 \cdot (-2) = -2$. 我们的 2×2 行列式公式用于

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

得 $1 \cdot 4 - 2 \cdot 3 = 4 - 6 = -2$.

若原来的矩阵 $\{a_{ij}\}$ 是非奇异的, 使得 $k = m$ 及 $n_i = i (i = 1, \cdots, n)$, 则在半约化矩阵中的第 j 列为 $c_j \delta^j$, 所以 $s_{jj} = c_j$, 及我们宣称行列式为首项系数的乘积 $\prod_{i=1}^m c_i$.

欲证这一点, 注意若 T 是对应于我们的半约化矩阵的 $\mathrm{Hom}(\mathbb{R}^n)$ 中的变换, 则 $T(\delta^j) = c_j \delta^j$, 所以 $\mathbb{R}^n$ 为 n 个 T - 不变量之直和, 它们是一维子空间, 在第 j 个上面, T 为 c_j 的乘积. 因此由我们列举的行列式性质 (c) 与 (d) 可知 $\Delta(T) = \prod_1^n c_j = \prod_1^n s_{jj}$. 它不为零.

另一方面, 若 $\{a_{ij}\}$ 为奇异的, 所以 $k = d(V) < m$, 则半约化矩阵中的第 m 行为 0, 特别 $s_{mm} = 0$. 因此乘积 $\prod_i s_{ii}$ 为零. 现在不改变主对角线, 我们可以减去含有首项行元素 (含有指标 n_j 之列) 之列的倍数使第 m 列为一个零列. 这个方法等价于右乘一个型 (2) 的初等矩阵, 所以行列式亦是不变的. 但是现在这个矩阵的变换 S 使 $\mathbb{R}^{m-1}$ 不变 (作为 $\mathbb{R}^m$ 中, $\delta^1, \cdots, \delta^{m-1}$ 的生成) 并取 δ^m 为 0, 所以根据列举的行列式性质 (c) 可知 $\Delta(S) = 0$. 因此行列式亦等于半约化矩阵的主对角线元素之乘积, 对这一情况为零.

我们亦得到一个矩阵为非奇异的 (可逆的) 当且仅当它的行列式非零.

习题

6.1 试计算出

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 2 & 3 & 2 \\ -1 & -3 & 0 & 4 \\ 0 & 4 & -1 & -3 \end{bmatrix}$$

的行空间的标准基.

6.2 对下面矩阵做同样的事

$$\begin{bmatrix} 1 & 2 & 4 & 5 \\ 1 & 2 & 3 & 4 \\ -1 & -2 & 0 & 2 \end{bmatrix}.$$

6.3 取不同的第一次选择, 对上面矩阵做同样的事.

6.4 试用行约化算出

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 7 \end{bmatrix}$$

之逆. 用乘法验证一下你的结果.

6.5 试用行约化

$$\begin{bmatrix} 1 & 2 & 3 & y_1 \\ 2 & 3 & 4 & y_2 \\ 3 & 4 & 7 & y_3 \end{bmatrix}.$$

在行约化后的矩阵中, 第 4 列如何与上面习题里算出的

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 7 \end{bmatrix}$$

之逆相比较? 解释一下.

6.6 试用行约化检验 $\prec 1,1,1,1 \succ, \prec 1,2,3,4 \succ, \prec 0,1,0,1 \succ$ 与 $\prec 4,3,2,1 \succ$ 是否为线性独立的. 在这个验证中行约化运算之一部分是不必要的, 那是什么?

6.7 若 $k \times n$ 矩阵 $\boldsymbol{a}$, 它以 α_i 为其第 i 行 $(1 \leqslant i \leqslant k)$, 是行约化梯形式, 则我们称 $\mathbb{R}^n$ 中向量的 k 元组 $\{\alpha_i\}_1^k$ 为标准的. 假定一个 n 元组 ξ 在 $\boldsymbol{a}$ 的行空间中, 我们可以读出关于上述标准基, 它的坐标. 它们是什么? 然后, 我们怎样才能验证一个任意的 n 元组 ξ 在行空间中与否?

6.8 试利用上面习题建议的行约化方法来决定 $\delta^1 = \prec 1,0,0,0 \succ$ 是否在 $\prec 1,1,1,1 \succ$, $\prec 1,2,3,4 \succ$ 与 $\prec 2,0,1,-1 \succ$ 的生成之中. 对 $\prec 1,2,1,2 \succ$ 与 $\prec 1,1,0,4 \succ$ 做同样的事.

6.9 假定 $a \neq 0$, 试由约化矩阵为梯形式来证明

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

为可逆的当且仅当 $ad - bc \neq 0$.

6.10 命 $\boldsymbol{a}$ 为一个 $m \times n$ 矩阵, 又命 $\boldsymbol{u}$ 为一个非奇异矩阵, 它行约化 $\boldsymbol{a}$, 所以 $\boldsymbol{r} = \boldsymbol{u} \cdot \boldsymbol{a}$ 为由 $\boldsymbol{a}$ 得到的行约化梯矩阵. 假定 $\boldsymbol{r}$ 有 $m - k > 0$ 个零行位于矩阵之底部 (第 k 行非零). 试证 $\boldsymbol{u}$ 的底部 $m - k$ 行生成 A 的值域的零化子 $(\text{range}A)^{\circ}$. 即对于某 $\boldsymbol{x}, \boldsymbol{y} = \boldsymbol{a}\boldsymbol{x}$ 当且仅当对于 $\boldsymbol{u}$ 的底部 $m - k$ 行中每一个 m 元组 $\boldsymbol{c}$ 有

$$\sum_1^m c_i y_i = 0.$$

[提示: $\boldsymbol{r}$ 的底部行是由 $\boldsymbol{u}$ 的底部行与 $\boldsymbol{a}$ 的列作用而得到的.]

6.11 回忆一下, 我们将约化 $\boldsymbol{a}$ 为 $\boldsymbol{r}$ 的行运算作用于 $m \times m$ 单位矩阵 $\boldsymbol{e}$ 来求得行约化矩阵 $\boldsymbol{u}$. 即我们行约化 $m \times (n+m)$ 并置矩阵 $\boldsymbol{a}|\boldsymbol{e}$ 至 $\boldsymbol{r}|\boldsymbol{u}$. 假定上面习题陈述的结果, 若 A 的矩阵为

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{bmatrix},$$

试求作为一个泛函的零空间, $A \in \text{Hom}(\mathbb{R}^3)$ 的值域.

6.12 类似地, 试求 A 的值域, 其中 A 的矩阵为

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \\ 4 & 2 & 3 \end{bmatrix}.$$

6.13 命 $\boldsymbol{a}$ 为一个 $m \times n$ 矩阵及 $\boldsymbol{a}$ 被行约化为 $\boldsymbol{r}$. 命 A 与 R 为在 $\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 中的对应算子 [所以 $A(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$]. 试证 A 与 R 有同样的零空间及 A^* 与 R^* 有同样的值空间.

6.14 试证解一个 n 个未知数的 m 个线性方程组等价于对于 n 元组 $\boldsymbol{x}$, 解一个矩阵方程

$$\boldsymbol{k} = \boldsymbol{t}\boldsymbol{x},$$

其中 $\boldsymbol{t}$ 为给予的 $m \times n$ 矩阵及 $\boldsymbol{k}$ 为给予的 m 元组. 命 $T \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 为被 $\boldsymbol{t}$ 乘. 从我们关于 T 的一般线性理论检查一下一个解的可能性 (值域, 零空间, 仿射子空间).

6.15 命 $\boldsymbol{b}=\boldsymbol{c}|\boldsymbol{d}$ 为并置 $m\times n$ 矩阵 $\boldsymbol{c}$ 与 $m\times p$ 矩阵 $\boldsymbol{d}$ 而得到的 $m\times(n+p)$ 矩阵. 若 $\boldsymbol{a}$ 为一个 $l\times m$ 矩阵, 试证

$$\boldsymbol{a}\cdot\boldsymbol{b}=\boldsymbol{ac}|\boldsymbol{ad}.$$

陈述关于 $\boldsymbol{b}$ 作为 n 列 m 元组的并置的类似表达式. 陈述在并置上右乘积的"分配性"的对应定理.

6.16 命 $\boldsymbol{a}$ 为一个 $m\times n$ 矩阵及 $\boldsymbol{k}$ 为一个列 m 元组. 命 $\boldsymbol{b}|\boldsymbol{1}$ 为 $m\times(n+1)$ 矩阵, 它是由 $m\times(n+1)$ 并置矩阵 $\boldsymbol{a}|\boldsymbol{k}$ 经行约化而得到的. 试证 $\boldsymbol{a}\cdot\boldsymbol{x}=\boldsymbol{k}$ 当且仅当 $\boldsymbol{b}\cdot\boldsymbol{x}=\boldsymbol{1}$. 再证存在一个解 $\boldsymbol{x}$ 当且仅当 $\boldsymbol{b}$ 中每一个为零的行, 在 $\boldsymbol{1}$ 中亦为零. 试用行秩的概念重述这个条件.

6.17 命 $\boldsymbol{b}$ 为由一个 $m\times n$ 矩阵 $\boldsymbol{a}$ 得到的行约化梯矩阵. 因此 $\boldsymbol{b}=\boldsymbol{u}\cdot\boldsymbol{a}$, 此处 $\boldsymbol{u}$ 为非奇异的, 而且 B 与 A 有同样的零空间 (此处 $B\in\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$ 为由 $\boldsymbol{b}$ 之乘积). 我们可以由 $\boldsymbol{b}$ 读出一个子空间 $W\subset\mathbb{R}^n$ 的一个基满足 $B\restriction W$ 是一个至值域 B 的同构. 这个基是什么? 然后我们得知 B 的零空间 N 是 W 的一个补. W 的一个补, 称它为 M, 它可以由 W 读出. M 是什么?

6.18 继续上面的习题, 试证对于 M 中每一个标准基向量 δ_i, 我们可以由矩阵 $\boldsymbol{b}$ 读出 W 中一个向量 α_i 适合 $\delta^i-\alpha_i\in N$. 再证这些向量 $\{\delta^i-\alpha_i\}$ 构成 N 的一个基.

6.19 假定第 5 节的性质 (a)~(e), 我们仍然要证明修改后的初等行运算使一个方阵的行列式不变. 首先, 由 (a),(c),(d) 与 (e) 证明若 $T\in\mathrm{Hom}\mathbb{R}^2$ 由 $T(\delta^1)=\delta^2$ 及 $T(\delta^2)=-\delta^1$ 定义, 则 $\Delta(T)=1$. 用一个非常简单的因子分解 $T=R\circ S$ 来做这件事, 此处 (e) 可以被用于 S. 从而一个 $(1')$ 型初等矩阵有行列式 1.

6.20 试由行列式的性质 (b) 证明一个 (2) 型的初等矩阵有行列式 1. 因此可证修改过的初等行运算作用于方阵上使它的行列式不改变.

*2.7 二次型的对角化

正如我们早先提到过的, 线性代数的中心问题之一为 $\mathrm{Hom}V$ 中一个线性变换 T 的"结构"的分析. 由基的观点, 这个领域的每一条定理皆断言, 选择 V 的一个特殊基, T 的矩阵可以被这种那种简单的形式给出. 这是这个主题非常难的部分, 尽管第一章定理 5.5 及其推论构成结构性结果的一块基石, 但在这本书中, 我们仅对这个问题作一点接触.

在本节我们仅去解决一个较简单的问题. 用上述语言就是这样的问题, 即选择 V 的一个基使在 $\mathrm{Hom}(V,V^*)$ 中变换 T 的矩阵简单化. 这样一个变换等价于

V 上一个双线性泛函 (由第一章的定理 6.1 与本章定理 3.2); 我们将按这个框架来处理问题.

命 V 为一个有限维实向量空间, 又命 $\omega: V \times V \to \mathbb{R}$ 为一个双线性泛函. 若 $\{\alpha_i\}_1^n$ 为 V 的一个基, 则 ω 决定了一个矩阵 $t_{ij} = \omega(\alpha_i, \alpha_j)$. 我们知道若 $\omega_\eta(\xi) = \omega(\xi, \eta)$, 则 $\omega_\eta \in V^*$ 及 $\eta \mapsto \omega_\eta$ 为一个由 V 至 V^* 的线性映射 T. 我们将给读者留作一个习题证明 $\{t_{ij}\}$ 是 T 关于 V 的基 $\{\alpha_i\}$ 及其关于 V^* 的对偶基的矩阵 (习题 4.1).

若 $\xi = \sum_1^n x_i\alpha_i$, 与 $\eta = \sum_1^n y_j\alpha_j$, 则

$$\omega(\xi, \eta) = \sum_{i,j} x_i y_j \omega(\alpha_i, \alpha_j) = \sum_{i,j} t_{ij} x_i y_j.$$

特别, 若我们置 $q(\xi) = \omega(\xi, \xi)$, 则 $q(\xi) = \sum_{i,j} t_{ij} x_i x_j$ 为一个坐标 x_i 的齐次二次多项式.

在本节的其余部分, 我们假定 ω 是对称的: $\omega(\xi, \eta) = \omega(\eta, \xi)$. 则我们可以从二次型 q, 经

$$\omega(\xi, \eta) = \frac{q(\xi + \eta) - q(\xi - \eta)}{4}$$

恢复 ω, 读者易于自己来验证. 特别若双线性型 ω 不恒等零, 则存在向量 ξ 满足 $q(\xi) = \omega(\xi, \xi) \neq 0$.

我们需要做的是去证明我们可以找到 V 的一个基 $\{\alpha_i\}_1^n$ 满足当 $i \neq j$ 时, $\omega(\alpha_i, \alpha_j) = 0$ 及 $\omega(\alpha_i, \alpha_i)$ 取三个值 $0, \pm 1$ 之一. 借助于标量乘积理论的标准用法 (见第五章), 我们称这样一个基为*标准正交的*. 我们关于一个标准正交基存在的证明将用 $n = \dim V$ 上的一个归纳法. 若 $n = 1$, 则任意非零向量 β 都是一个基, 又若 $\omega(\beta, \beta) \neq 0$, 则我们可以选择 $\alpha = x\beta$ 使 $x^2\omega(\beta, \beta) = \omega(\alpha, \alpha) = \pm 1$, 其中 x 需取之值显然为 $x = |\omega(\beta, \beta)|^{-1/2}$. 在一般情况, 若 ω 为一个零泛函, 则任意基将显然地为标准正交的, 因此我们可以假定 ω 不恒等于 0. 则恰如我们在较早时说的, 存在一个 β 使 $w(\beta, \beta) \neq 0$. 我们置 $\alpha_n = x\beta$, 此处 x 被选取使 $q(\alpha_n) = \omega(\alpha_n, \alpha_n) = \pm 1$. 这个非零线性泛函 $f(\xi) = \omega(\xi, \alpha_n)$ 有一个 $(n-1)$ 维零空间 N, 及若我们命 ω' 为 ω 在 $N \times N$ 上的限制, 则由归纳法假定可知 ω' 有一个标准正交基 $\{\alpha_i\}_1^{n-1}$. 同时, 因为当 $i < n$ 时, α_i 在 f 的零空间之中, 所以 $\omega(\alpha_i, \alpha_n) = \omega(\alpha_n, \alpha_i) = 0$. 因此 $\{\alpha_i\}_1^n$ 是 ω 的一个标准正交基, 于是我们达到了目的:

定理 7.1 若 ω 为一个有限维实向量空间 V 上的一个对称双线性泛函, 则 V 有一个 ω 标准正交基.

对于一个 ω 标准正交基, 表达式 $\omega(\xi,\eta)=\sum x_i y_j \omega(\alpha_i,\alpha_j)$ 归化为

$$\omega(\xi,\eta)=\sum_{i=1}^{n} x_i y_i q(\alpha_i),$$

此处 $q(\alpha_i)=\pm 1$ 或 0. 若我们命 V_1 为使 $q(\alpha_i)=1$ 的基向量 α_i 的生成, 及类似地定义 V_{-1} 与 V_0, 则我们得知对于每一个 V_1 中的非零 ξ 有 $q(\xi)>0$; 对于每一个 V_{-1} 中的非零 ξ 有 $q(\xi)<0$, 及在 V_0 上有 $q=0$. 而且, $V=V_1 \bigoplus V_{-1} \bigoplus V_0$ 及三个子空间彼此是 ω 标准正交的 (意思是若 $\xi \in V_1$ 及 $\eta \in V_{-1}$, 则 $\omega(\xi,\eta)=0$ 等等). 最后, 对于每一个 $\xi \in V_{-1} \bigoplus V_0$ 皆有 $q(\xi) \leqslant 0$.

若我们取另一个标准正交基 $\{\beta_i\}$ 而命 W_1, W_{-1} 与 W_0 为它的对应子空间, 则 W_1 可能与 V_1 不一样, 但它们的维数必须相同. 我们有 $W_1 \cap (V_{-1} \bigoplus V_0)=\{0\}$, 这是由于在这个交中的任意非零 ξ 将导出矛盾的不等式 $q(\xi)>0$ 与 $q(\xi) \leqslant 0$. 因此 W_1 可以被延拓为 $V_{-1} \bigoplus V_0$ 的一个补, 及由于 V_1 是一个补, 所以 $d(W_1) \leqslant d(V_1)$. 类似地, $d(V_1) \leqslant d(W_1)$, 所以它们的维数相等. 附带地, 这证明了 W_1 是 $V_{-1} \bigoplus V_0$ 的一个补. 用完全相同的方法, 我们求得 $d(W_{-1})=d(V_{-1})$, 于是最后, 用减法即得到 $d(W_0)=d(V_0)$. 按照惯例, ω 标准正交基 $\{\alpha_i\}_1^n$ 可以重新排序使满足 $q(\alpha_i)=1$ 的所有 α_i 排在最前面, 然后是满足 $q(\alpha_i)=-1$, 及最后为满足 $q(\alpha_i)=0$. 我们将上述结果陈述如下:

定理 7.2 若 ω 是一个有限维空间 V 上的一个对称双线性泛函, 则存在整数 n 与 p 使若 $\{\alpha_i\}_1^m$ 为任意按惯例次序的 ω 标准正交基, 及若 $\xi=\sum_1^m x_i\alpha_i$, 则

$$q(\xi)=x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_{p+n}^2=\sum_1^p x_i^2-\sum_{p+1}^{p+n} x_i^2.$$

整数 $p-n$ 称为 q 型 (或它对应的对称双线性泛函 ω) 的符号差,而 $p+n$ 为它的秩. 注意 $q+n$ 是 q 的上述矩阵的列空间的维数, 所以它等于相关线性映射 T 的值域之维数. 因此 $p+n$ 为 q 的每一个矩阵的秩.

一个标准正交基的归纳法证明并不表示我们在实际上如何找出它来. 假定我们前面有关于某基 $\{\alpha_i\}_1^n$, ω 的矩阵 $\{t_{ij}\}$, 使得 $\omega(\xi,\eta)=\sum x_i y_j t_{ij}$, 此处 $\xi=\sum_1^n x_i\alpha_i$, $\eta=\sum_1^n y_i\alpha_i$, 及 $t_{ij}=\omega(\alpha_i,\alpha_j)$, 因而我们需要知道如何真正地求出一个标准正交基 $\{\beta_i\}_1^n$. 主要问题是找到一个正交基; 标准化则是平凡的. 第

一件事为找到一个向量 β 满足 $\omega(\beta,\beta)\neq 0$. 若某 $t_{ii}=\omega(\alpha_i,\alpha_i)\neq 0$, 则我们可以取 $\beta=\alpha_i$. 若所有 $t_{ii}=0$, 而 ω 不是零型, 则必定有某 $t_{ij}\neq 0$, 假定 $t_{12}\neq 0$. 若我们置 $\gamma_1=\alpha_1+\alpha_2$ 及 $\gamma_i=\alpha_i(i>1)$, 则 $\{\gamma_i\}_1^n$ 为一个基, 及 ω 关于基 $\{\gamma_i\}$ 的矩阵 $s=\{s_{ij}\}$ 有

$$s_{11}=\omega(\gamma_1,\gamma_1)=\omega(\alpha_1+\alpha_2,\alpha_1+\alpha_2)=t_{11}+2t_{12}+t_{22}=2t_{12}\neq 0.$$

类似地, 若 i 或 j 大于 1 时, $s_{ij}=t_{ij}$.

例如, 若 ω 为由 $\omega(\boldsymbol{x},\boldsymbol{y})=x_1y_2+x_2y_1$ 定义的 $\mathbb{R}^n$ 上的双线性型, 则它的矩阵 $t_{ij}=\omega(\delta^i,\delta^j)$ 为

$$\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},$$

于我们必须换基使 $t_{11}\neq 0$. 按照上面的方式, 我们置 $\gamma_1=\delta^1+\delta^2$ 及 $\gamma_2=\delta^2$, 从而得到新矩阵 $s_{ij}=\omega(\gamma_i,\gamma_j)$, 即

$$\begin{bmatrix}2 & 1\\ 1 & 0\end{bmatrix},$$

下一步为寻找泛函 $\omega(\xi,\gamma_1)=\sum x_is_{1i}$ 的零空间的一个基. 我们用修改 $\gamma_2,\cdots,\gamma_n$ 来做这件事. 我们将 γ_j 换成 $\gamma_j+c\gamma_1$ 及算出 c 使这个向量在零空间之中. 因此我们需要 $0=\omega(\gamma_j+c\gamma_1,\gamma_1)=s_{1j}+cs_{11}$, 从而 $c=-s_{1j}/s_{11}$. 注意在我们做到 $s_{11}\neq 0$ 之前, 我们不能采取这个正交化步骤. 这个新的集合仍为生成, 所以它是一个基, 而新的矩阵 $\{r_{ij}\}$ 有 $r_{11}\neq 0$ 及当 $j>1$ 时, $r_{1j}=r_{j1}=0$. 我们现在仅仅对 ω 在这个 $(n-1)$ 维零空间的限制来重复整个的程序, 其矩阵为 $\{r_{ij}:2\leqslant i,j\leqslant n\}$ 等等. 这是长的过程, 但是到标准化, 我们只在原来矩阵上用了有理运算, 我们用加, 减, 乘与除, 但我们没有去找多项式方程的根.

继续我们上面的例子, 我们置 $\beta_1=\gamma_1$, 但我们将 γ_2 换成 $\beta_2=\gamma_2-(s_{12}/s_{11})\cdot\gamma_1=\gamma_2-1/2\gamma_1$. 最终的矩阵 $r_{ij}=\omega(\beta_i,\beta_j)$ 有

$$r_{11}=s_{11}=2,\quad r_{12}=\omega(\beta_1,\beta_2)=\omega(\gamma_1,\gamma_2-\frac{1}{2}\gamma_1)=s_{12}-\frac{1}{2}s_{11}=0,$$

及 $r_{22}=\omega(\gamma_2-\frac{1}{2}\gamma_1,\gamma_2-\frac{1}{2}\gamma_1)=s_{22}-s_{12}+s_{11}/4=-\frac{1}{2}$, 所以

$$\{r_{ij}\}=\begin{bmatrix}2 & 0\\ 0 & -1/2\end{bmatrix},$$

最终的基为 $\beta_1 = \gamma_1 = \delta^1 + \delta^2$ 及 $\beta_2 = \gamma_2 - \frac{1}{2}\gamma_1 = \delta^2 - \frac{1}{2}(\delta^1 + \delta^2) = (\delta^2 - \delta^1)/2$.

我们在上面采取的步骤为行约化之回顾, 但是由于我们对具有 ω 的变换 $T: V \to V^*$ 的定义空间与值空间同时变换基, 所以每一步骤都同时涉及一个初等矩阵的左乘与右乘. 即我们同时作行与列的约化. 这在直观上应该是清楚的, 即按这样的方式作用于一个对称矩阵, 可使它保持对称.

关于二次型另外的信息, 我们回到关于一个变换的矩阵的基变换公式:$t'' = \boldsymbol{b} \cdot \boldsymbol{t}' \cdot \boldsymbol{a}$. 在此 T 是 ω 型由 V 至 V^* 的变换, 所以按照 §4 的计算有 $\boldsymbol{b} = (\boldsymbol{a}^*)^{-1}$. 行列式函数的性质之一为 $\Delta(T^*) = \Delta(T)$, 所以 $\Delta(\boldsymbol{a}^*) = \Delta(\boldsymbol{a})$. 因此若 $\boldsymbol{s}$ 与 $\boldsymbol{t}$ 为一个二次型关于 V 的第一个与第二个基的矩阵. 又若 $\boldsymbol{a}$ 是基变换矩阵, 则 $\boldsymbol{s} = (\boldsymbol{a}^*)^{-1} \cdot \boldsymbol{t} \cdot \boldsymbol{a}^{-1}$ 及 $\Delta(\boldsymbol{s}) = (\Delta(\boldsymbol{a}^{-1}))^2 \Delta(\boldsymbol{t})$. 因此一个二次型有*奇偶性*. 若它是非奇异的, 则它的行列式或者总是正的, 或者总是负的, 我们可以称它为偶或奇的. 在我们的例子里, 开始与终结的矩阵为

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad 与 \quad \begin{bmatrix} 2 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix},$$

它们的行列式都等于 -1.

在二维情况, 关于一个标准正交基的一个型的行列式, 当对角线上元素均为 $+1$ 或均为 -1 时, 等于 1; 当它们取不同符号时, 行列式为 -1. 因此*不作标准正交化*, 我们就可以读出二维空间一个非奇异型的符号差. 若行列式 $t_{11}t_{22} - (t_{12})^2$ 为正的, 则符号差为 ± 2, 我们由观察 t_{11} 即可断定是哪一个 (因在我们的正交化程序中, t_{11} 是不变的). 因此符号差为 $+2$ 或 -2 依赖于 $t_{11} > 0$ 或 $t_{11} < 0$. 若行列式为负的, 则符号差为 0. 因此不作计算, 我们就知道具有矩阵

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

的型 $\omega(\boldsymbol{x}, \boldsymbol{y}) = x_1y_2 + x_2y_1$ 的符号差为 0.

定理 7.1 与 7.2 对于向量空间上实值函数的临界点分类是很重要的. 我们在 3.16 节将看到这样一个函数 F 的二阶微分是一个对称的双线性泛函, 它的型的符号差与下面事情有同等重要性, 即决定 F 在一点近旁的性质, 在该点它的一阶导数为零, 在初等微积分中考虑了二阶导数的符号.

若除去 $\xi = 0$ 外, $q(\xi)$ 皆非零, 则称二次型 q 为*定的*. 所以 $q(\xi)$ 必须总有相同的符号, 这样将 q 称为*正定的或负定的*. 回过来看看定理 7.2, 这应该是显然的, 即仅当 $p = d(V)$ 及 $n = 0$ 时, q 才是正定的, 又仅当 $n = d(V)$ 及 $p = 0$ 时, q

才是负定的. 一个对称的双线性泛函, 其相关的二次型是正定的则称为一个标量积. 在一般向量空间中, 这是一个非常重要的概念, 整个第五章都将从事于发展它的一些内含.

第三章　微分学

现在我们的代数背景对微分学来说已经合适了，然而我们还需要一些多维的极限理论。粗略地说，微分学是对非线性映射的线性近似理论，从而我们不得不了解在一般的向量架构下近似到底意味着什么。因此我们将从研讨一个向量空间 V 中向量的长度度量概念作为本章的开端，这里的向量长度度量被称作范数.然后我们便能研究由一个曲面的切平面去近似在切点附近的此曲面的这种方式所产生的现象. 这是被称为微分的关于映射的惟一局部线性近似的一般理论. 计算微分的规则中包括了微分学中所有熟悉的规则，并且达到了使复杂计算可按常规方式运作的同一目的. 但是在多维情况下这个理论更为丰富，而且我们必须要掌握一个新的东西，即线性变换和它们在给定向量上取值之间的相互关联，前者是微分而后者在一般情形下是方向导数，在向量属于某组基时则为偏导数. 特别地，当所考虑的空间为有限维并通过选取基将其置换成笛卡儿空间时，此微分则完全等价于它的矩阵，它是一个由偏导数构成的矩阵并被称作此映射的雅可比矩阵. 于是微分运算的规则便用矩阵运算的方式表达出来.

求实函数的极大和极小点完全像从前那样由计算微分并使它等于零而得到. 但是我们将略去这个课题，只在带星号的章节中有所讨论. 它也远较其单变的相应部分更为丰富，并且在某种无穷维情形中成为被称作变分法的学科.

最后，我们将开始我们对逆映射定理和隐函数定理的研讨. 逆映射定理说，如果向量空间之间的一个映射是连续可微的，并且如果其微分在某点 α 是可逆

的 (作为一个线性变换), 则此映射本身在 α 的某个邻域内是可逆的. 隐函数定理说, 如果两个向量变量的一个向量值函数 G 被令为零, 并且如果 G 的第二个偏微分(作为线性映射) 在点 $\prec \alpha, \beta \succ$ 可逆, 其中 $G(\alpha, \beta) = 0$, 则方程 $G(\xi, \eta) = 0$ 可以在此点附近解出用 ξ 表达的 η. 就是说, 存在一个在 α 附近定义的惟一决定的映射 $\eta = F(\xi)$ 使得 $\beta = F(\alpha)$, 并且使得在 α 的此邻域中有 $G(\xi, F(\xi)) = 0$.

这两个定理对分析学的进一步展开是基础性的. 较之于迄今为止的工作它们之所以是更为深刻的在于它们依赖于向量空间的一个特殊性质, 即*完备性*; 我们将不得不把它们的部分证明推延到下一章, 在那里我们要以完全系统的方式研究完备性.

在本章末尾的许多带星号的小节中我们提供了一些较难的素材, 我们并不期望读者能掌握它们. 但是他应该了解这些素材的大体概念.

3.1 回顾 ℝ 中的情形

假定每个学过微积分的学生都熟悉实数系的性质和极限理论. 但是在此我们只是熟悉还不够. 同学们理解 ε - 定义并能够使用它们是绝对必需的.

为稳妥起见, 我们将在函数极限的架构下回顾部分这些内容; 自信的读者可略去不读. 我们假定所有我们考虑的函数至少在包含 a 的一个开区间上有定义, 但可以除去 a 本身. 这个例外情形之所以需要是由计算中的差商表明的, 这时的差商在此点没有定义而它们在此点附近的行为才是至关重要的.

定义 1.1 *当 x 趋向于 a 时 $f(x)$ 趋向于 l(用符号表示即当 $x \to a$ 时 $f(x) \to l$) 是说, 对每个正数 ε 存在一个正数 δ 使得*

$$0 < |x - a| < \delta \Rightarrow |f(x) - l| < \varepsilon.$$

我们也说当 x 趋向 a 时, l 是 $f(x)$ 的极限, 写为 $\lim_{x\to a} f(x) = l$. 在此定义中所给出的陈述可以理解为对 x 的全称量化, 故而此定义真正开始于三个量词 $(\forall\varepsilon^{>0})(\exists\delta^{>0})(\forall x)$. 这些前置的量词使得定义被当作文章阅读时显得造作和不符合词言习惯, 但是读者应记住在我们对量化的最初讨论中所表示的, 即为了使语句的意义清晰而无任何含糊之处, 这种人为制做的语句是绝对必要的. *对量词 $(\forall\varepsilon)(\exists\delta)(\forall x)$ 顺序的任何改变将改变此陈述的意义*

内置全称量词限定语句

$$(\forall x)(0 < |x - a| < \delta \Rightarrow |f(x) - l| < \varepsilon)$$

的意义是直观的并容易以图形表示 (见图 3.1).

对所有比 δ 更靠近 a 的 x, f 在 x 的值比 ε 更靠近 l. 此定义一开始所说的是, 对每个正数的 ε 可以找到这样的一个正数 δ. 自然, δ 随 ε 而变化; 如果 ε 给得越小, 一般地我们则必定走向 a 越近, 那就是说, 使在 $(a-\delta, a+\delta)-\{a\}$ 上 f 的值以 ε 靠近 l 之前, 我们必须把 δ 取得较小.

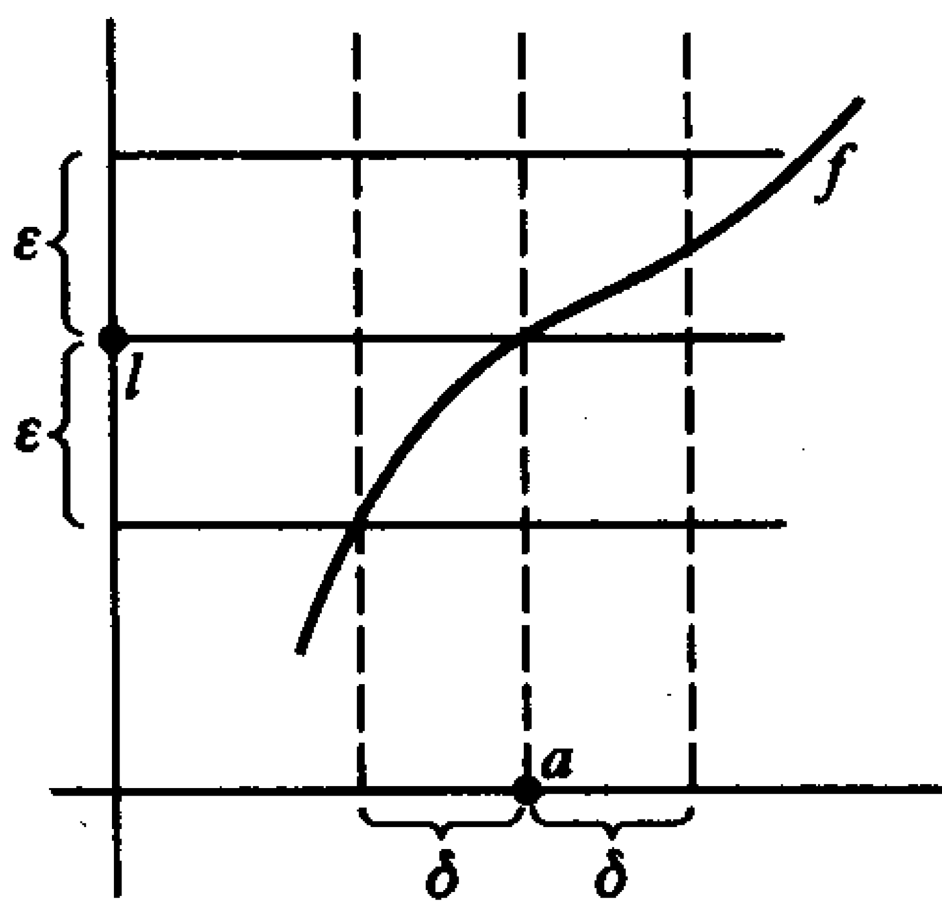


图 3.1

变量 'ε' 和 'δ' 几乎总被限制为正实数, 故从现在起我们让此限制只隐含而不明写出来, 除非有某些特殊要求. 因此我们将简单地写成 $(\forall\varepsilon)(\exists\delta)\cdots$.

收敛性的定义以各种方式被运用. 在最简单的情形中, 假设我们已知一个或更多的在 a 有极限的函数, 譬如, $f(x)\to u$, 和 $g(x)\to v$, 而我们要证明某个其他的函数 h 在 a 有极限 w. 在这种情形中, 我们总是试图找到一个不等式, 以我们已知可以做得很小的量 $|f(x)-u|$ 和 $|g(x)-v|$ 来表达我们想要使它变小的量 $|h(x)-w|$.

例如, 假定 $h=f+g$. 由于 $f(x)$ 靠近 u, $g(x)$ 靠近 v, 则显然 $h(x)$ 靠近 $w=u+v$. 但是如何靠近? 因为 $h(x)-w=(f(x)-u)+(g(x)-v)$, 我们有

$$|h(x)-w|\leqslant|f(x)-u|+|g(x)-v|.$$

由此我们清楚, 为了使 $|h(x)-w|$ 小于 ε 只要使 $|f(x)-u|$ 和 $|g(x)-v|$ 中每一个都小于 $\varepsilon/2$ 即可. 因此, 给定任意 ε, 我们可取 δ_1, 使得 $0<|x-a|<\delta_1\Rightarrow|f(x)-u|<\varepsilon/2$, 以及 δ_2, 使得 $0<|x-a|<\delta_2\Rightarrow|g(x)-v|<\varepsilon/2$, 于是我们可以取 δ 为这两个数中较小者, 故而如果 $0<|x-a|<\delta$, 则两个不等式都成立. 因而

$$0<|x-a|<\delta\Rightarrow|h(x)-w|\leqslant|f(x)-u|+|g(x)-v|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

那么, 我们对函数 h 便找到了所想要的 δ.

下面假定 $u\neq0$, 而 $h=1/f$. 显然当 $f(x)$ 靠近 u 时 $h(x)$ 则靠近 $w=1/u$, 因而我们想把 $h(x)-w$ 用 $f(x)-u$ 来表达. 因此

$$h(x)-w=\frac{1}{f(x)}-\frac{1}{u}=\frac{u-f(x)}{f(x)u},$$

从而 $|h(x)-w|\leqslant|f(x)-u|/|f(x)u|$. 这里的麻烦在于分母是变量, 而且如果它碰巧非常小则会抵消了 $|f(x)-u|$ 的小, 从而不能迫使此商很小. 然而容易解

决这个问题. 由于 $f(x)$ 靠近 u 而 u 又非零, 故 $f(x)$ 不会靠近零. 例如, 如果 $f(x)$ 比 $u/2$ 更靠近 u, 则 $f(x)$ 离开 0 必比 $|u|/2$ 更远. 因此我们选取 δ_1 使得 $0<|x-a|<\delta_1 \Rightarrow |f(x)-u|<|u|/2$, 由此得到 $|f(x)|>|u|/2$. 于是

$$|h(x)-w|<2|f(x)-u|/|u|^2,$$

那么, 给定任一 ε, 我们选取 δ_2 使得

$$0<|x-a|<\delta_2 \Rightarrow |f(x)-u|<\varepsilon|u|^2/2.$$

再取 δ 为 δ_1 和 δ_2 中较小者, 故当 $0<|x-a|<$时这两个不等式都成立, 我们有

$$0<|x-a|<\delta \Rightarrow |h(x)-w|<2|f(x)-u|/|u|^2<2\varepsilon|u|^2/2|u|^2=\varepsilon,$$

我们又对函数 h 找到了我们的 δ.

我们已力图表明人们是如何考虑这些情况的. 写下来的实际证明只须指出 δ 的选取. 因此,

引理 1.1 如果当 $x\to a$ 时 $f(x)\to u$, $g(x)\to v$, 则当 $x\to a$ 时 $f(x)+g(x)\to u+v$.

证明 给出 ε, 选取 δ_1 使得 $0<|x-a|<\delta_1 \Rightarrow |f(x)-u|<\varepsilon/2$(因为已假定了 f 在 a 对 u 的收敛性), 相似地, 选取 δ_2 使得 $0<|x-a|<\delta_2 \Rightarrow |g(x)-v|<\varepsilon/2$. 取 δ 为 δ_1 和 δ_2 中的较小者. 于是

$$0<|x-a|<\delta \Rightarrow |(f(x)+g(x))-(u+v)|\leqslant|f(x)-u|+|g(x)-v|<\varepsilon/2+\varepsilon/2=\varepsilon.$$

从而得证. □

除了了解极限理论的 ε - 技术外, 还有必要理解并能够运用实数系的基本性质即上确界性质. 在下面对此性质的陈述中的半无穷区间 $(-\infty,a]$ 自然指的是子集 $\{x\in\mathbb{R}:x\leqslant a\}$.

如果 A 是 $\mathbb{R}$ 的一个非空子集, 使得对一个数 a 有 $A\subset(-\infty,a]$ 则存在一个惟一确定的最小数 b, 使得 $A\subset(-\infty,b]$.

称满足 $A\subset(-\infty,a]$ 的数 a 为 A 的一个上界; 显然, a 是 A 的一个上界当且仅当 A 中的每个 x 均小于或等于 a, 有一个上界的集合被叫做上有界的. 上面的性质说一个上有界的非空集合 A 有一个最小上界即上确界 (lub). 如果用 -1 乘以每个数则逆转了顺序关系, 于是我们得到了另一个陈述, 它断言 $\mathbb{R}$ 的任一下有

界的非空子集合有一个最大下界即下确界(glb). 区间 (0,1) 的上确界为 1. [0,1] 的上确界也是 1. $\{1/n : n$ 为正整数$\}$的下确界为 0. 还有, lub$\{x : x$ 为正有理数且 $x^2 < 2\} = \sqrt{2}$, glb$\{e^x : x \in \mathbb{R}\} = 0$, 以及 lub$\{e^x : x$ 为有理数且 $x < \sqrt{2}\} = e^{\sqrt{2}}$.

习题

1.1 证明如果当 $x \to a$ 时 $f(x) \to l$ 和 $f(x) \to m$, 则 $l = m$. 因此我们可以谈及当 $x \to a$ 时 $f(x)$ 的惟一确定的极限.

1.2 证明如果当 $x \to a$ 时 $f(x) \to l$, $g(x) \to m$, 则当 $x \to a$ 时 $f(x)g(x) \to lm$.

1.3 证明 $|x - a| \leqslant |a|/2 \Rightarrow |x| \geqslant |a|/2$.

1.4 由下确界性质证明 (详细地) 上确界性质.

1.5 试证 lub$A = x$ 当且仅当 x 是 A 的一个上界, 而且对每个正的 ε, $x - \varepsilon$ 不是 A 的上界.

1.6 设 A 和 B 为 $\mathbb{R}$ 的子集, 它们均非空且上有界. 证明 $A + B$ 为非空, 上有界且 $\text{lub}(A + B) = \text{lub}A + \text{lub}B$.

1.7 准确阐述并证明关于两个集合乘积的上确界的一个正确定理.

1.8 对定义域为 $\mathbb{R}$ 的子集的函数定义单边极限的概念. 例如, 我们需要当 x 从下面趋向于 a 时 $f(x)$ 的极限, 我们记这种极限为

$$\lim_{x \uparrow a} f(x).$$

1.9 如果实函数 f 的定义域是一个区间, 譬如 $[a, b]$, 如果

$$x < y \Rightarrow f(x) \leqslant f(y)$$

我们则称 f 是一个增函数. 证明一个增函数处处有单边极限.

1.10 设 $[a, b]$ 为 $\mathbb{R}$ 中的闭区间, 并设 $f : [a, b] \to \mathbb{R}$ 为增函数. 证明 $\lim_{x \to y} f(x) = f(y)$ 对所有 $[a, b]$ 中 y 成立 (即 f 在 $[a, b]$ 上连续) 当且仅当 f 的值域没有漏掉任何一个子区间 $(c, d) \subset [f(a), f(b)]$.[提示: 假如此值域漏掉了 (c, d), 并令 $y = \text{lub}(x : f(x) \leqslant c)$. 于是当 $x \to y$ 时 $f(x) \nrightarrow f(y)$.]

1.11 一个集合与区间 $[s, t]$ 中每个开子集都相交则说此集合在 $[s, t]$ 中稠密. 证明如果 $f : [a, b] \to \mathbb{R}$ 为增函数且 f 的值域在 $[f(a), f(b)]$ 中稠密, 则 f 的值域 $=[f(a), f(b)]$.(对 $f(a)$ 和 $f(b)$ 之间的任意 z, 令 $y = \text{lub}\{x : f(x) \leqslant z\}$, 等等.)

1.12 假定上面两个习题的结果成立, 证明如果 f 是 $[a, b]$ 到 $\mathbb{R}$ 的连续的严格增函数, 并且如果 $r = f(a)$, $s = f(b)$, 则 f^{-1} 是从 $[r, s]$ 到 $\mathbb{R}$ 的连续的严格增函数.[如果对定义域

中每个 y 当 $x\to y$ 时 $f(x)\to f(y)$ 则函数 f 为连续; 如果 $x<y\Rightarrow f(x)<f(y)$ 则说 f 是严格增的.]

1.13 多少像上面习题 1.11 的论证那样证明, 如果 $f:[a,b]\to\mathbb{R}$ 在 $[a,b]$ 上连续, 则 f 的值域包含了 $[f(a),f(b)]$. 这是*介值定理*.

1.14 假设函数 $q:\mathbb{R}\to\mathbb{R}$ 满足对所有 $x,y\in\mathbb{R}$ 有 $q(xy)=q(x)q(y)$. 注意, $q(x)=x^n$(n 为正整数) 和 $q(x)=|x|^r$(r 为任意实数) 均满足此 "函数方程". $q(x)\equiv 0$ 也满足此方程 (可否理解为 $r=-\infty$?). 证明如果 q 满足此函数方程且对 $x>1$ 有 $q(x)>1$, 则存在一个实数 $r>1$ 使得对所有正数 x 有 $q(x)=x^r$.

1.15 证明如果 q 连续并对所有 $x,y\in\mathbb{R}$ 满足函数方程 $q(xy)=q(x)q(y)$, 并且至少存在一个点 a 使 $q(a)\neq 0,1$, 则对所有正数 x, $q(x)\equiv x^r$. 最后如果 q 也是非负的, 则在 $\mathbb{R}$ 上 $q(x)\equiv|x|^r$.

1.16 证明如果 $q(x)\equiv|x|^r$, 且如果 $q(x+y)\leqslant q(x)+q(y)$, 则 $r\leqslant 1$.(试一试 $y=1$ 和大的 x 值, 如果 $r>1$, $q'(x)$ 像是什么?)

3.2 范数

像上面所简短回顾的那样, 在 $\mathbb{R}$ 的极限理论中绝对值函数被显著地用于像 '$|x-y|$' 这样的表达式中, 以表示两个数之间的距离. 在这里就是 x 与 y 之间的距离. $f(x)$ 到 u 的收敛性定义只不过是下述意义的小心陈述, 即当距离 $|x-a|$ 趋向零时距离 $|f(x)-u|$ 趋向于零. 在我们证明中用到的 $|x|$ 的性质为

(1) $|x|>0$, 如果 $x\neq 0$, 且 $|0|=0$;

(2) $|xy|=|x||y|$;

(3) $|x+y|\leqslant|x|+|y|$.

向量空间的极限理论是用被称作*范数*的函数进行研究的, 它是 $\mathbb{R}$ 上绝对值函数在多维情形的类比. 因此, 如果 $p:V\to\mathbb{R}$ 是个范数, 我们则要把 $p(\alpha)$ 解释为 α 的 "大小", 而 $p(\alpha-\beta)$ 解释为 α 与 β 之间的 "距离". 但是如果 V 不是一维的, 则没有最自然的关于大小的概念. 例如, 如果 f 是 $[a,b]$ 上的正连续函数, 并且如果我们要求读者给出一个用于度量 f 是如何 "大" 的数, 对他而言大体会出现两种可能的选取: f 的最大值和 f 的图像下方的面积. 的确, 如果 $\max f$ 是小的则必定认为 f 为小. 但是如果 f 的那个面积为小, 我们也会在一种不同的意义下同意 f 也是小的. 在 $[a,b]$ 上连续的所有函数组成的向量空间 $V=\mathfrak{C}([a,b])$ 上有两个范数的例子:

$$p(f)=\max\{|f(t)|:t\in[a,b]\}$$

和

$$q(f)=\int_a^b |f(t)|dt.$$

请注意, f 可能在第二个意义是小的但是在第一个却不是的.

为了能有效利用, 向量的大小概念必须具有像 $\mathbb{R}$ 上绝对值函数相类比的性质.

定义 2.1 一个范数是向量空间 V 上的一个实函数, 满足

n1. $p(\alpha)>0$, 如果 $\alpha\neq 0$(正性);

n2. $p(x\alpha)=|x|p(\alpha)$, 对所有 $\alpha\in V$, $x\in\mathbb{R}$(齐性);

n3. $p(\alpha+\beta)\leqslant p(\alpha)+p(\beta)$, 对所有 $\alpha,\beta\in V$(三角不等式).

一个赋范线性空间(nls) 或者说赋范向量空间是一个向量空间 V 连同 V 上的一个范数. 赋范线性空间实则是一个偶对 $\prec V,p\succ$, 但是一般情况下我们只谈及赋范空间 V, 而理解这时在 V 上已有一个确定的范数.

按习惯用 $\|\alpha\|$ 表示 α 的范数, 大概是为了提示与绝对值的类比性. 于是三角不等式 n3 成为 $\|\alpha+\beta\|\leqslant\|\alpha\|+\|\beta\|$, 在形式上它几乎恒同于基本的绝对值不等式 $|x+y|\leqslant|x|+|y|$. 相似地, n2 成为 $\|x\alpha\|=|x|\|\alpha\|$. 类比于 $\mathbb{R}$ 中的 $|xy|=|x||y|$. 另外, $\|\alpha-\beta\|$ 可相似地解释为 α 与 β 之间的距离. 其合理性在于, 如果我们令 $\alpha=\xi-\eta,\beta=\eta-\zeta$, 则 n3 成为通常的几何的三角不等式:

$$\|\xi-\zeta\|\leqslant\|\xi-\eta\|+\|\eta-\zeta\|.$$

我们对范数将同时使用双重线记号和 “p” 记号; 有时其中一个比另一个更好.

在 $\mathbb{R}^n$ 上最通常使用的范数是 $\|\boldsymbol{x}\|_1=\sum_1^n|x_i|$, 欧氏范数 $\|\boldsymbol{x}\|_2=(\sum_1^n x_i^2)^{1/2}$, 和 $\|\boldsymbol{x}\|_\infty=\max\{|x_i|\}_i$. $[a,b]$ 上所有连续实函数的无穷维向量空间 $\mathfrak{C}([a,b])$ 上的范数相似地为

$$\|f\|_1=\int_a^b|f(t)|dt;$$

$$\|f\|_2=\left(\int_a^b|f(t)|^2dt\right)^{1/2};$$

$$\|f\|_\infty=\max\{|f(t)|:a\leqslant t\leqslant b\}.$$

对读者来说验证 $\|\ \|_1$ 在上面两种情形中均是个范数应该是容易的, 关于所谓的一致范数 $\|\ \|_\infty$ 我们将在下一节采用. 欧氏范数 $\|\ \|_2$ 则更富于技巧性; 它们

的性质依赖于所考虑的内积. 在第五章要讨论这些. 同时, 为使读者能在 $\mathbb{R}^n$ 上使用欧氏范数 $\| \|_2$, 我们要求他去证明它的三角不等式 (其他的公理是显然的), 这是一个以生硬方式去证明的习题. 在 $\mathbb{R}$ 自身, 绝对值便是一个范数, 它是仅有的在常数范围内确定的范数.

我们用下面的一般性评注将上面 $\mathbb{R}^n$ 的范数迁移到任意有限维空间上.

引理 2.1 如果 p 是向量空间 W 上的一个范数, T 是从向量空间 V 到 W 的单射的线性映射, 则 $p \circ T$ 是 V 上的一个范数.

证明 留给读者去证. □

一致范数 上面考虑的两个范数 $\| \|_\infty$ 是一个非常一般情形的特殊例子. 设 A 为任一非空集合, 并设 $\mathfrak{B}(A,\mathbb{R})$ 为所有有界函数 $f : A \to \mathbb{R}$ 的集合. 就是说, $f \in \mathfrak{B}(A,\mathbb{R})$ 当且仅当 $f \in \mathbb{R}^A$ 且对某个 $b \in \mathbb{R}$ 有 f 的值域 $\subset [-b,b]$. 这等于说 $|f|$ 的值域 $\subset [0,b]$, 并称这样的 b 为 $|f|$ 的一个界. 集合 $\mathfrak{B}(A,\mathbb{R})$ 是一个向量空间 V, 这是因为如果 $|f|$ 和 $|g|$ 分别以 b 和 c 为界, 则 $|xf + yg|$ 以 $|x|b + |y|c$ 为界. 定义一致范数 $\|f\|_\infty$ 为 $|f|$ 的最小界. 即

$$\|f\|_\infty = \text{lub}\{|f(p)| : p \in A\}.$$

当然, 必须验证 $\| \|_\infty$ 是个范数, 对任意 $p \in A$,

$$|f(p) + g(p)| \leqslant |f(p)| + |g(p)| \leqslant \|f\|_\infty + \|g\|_\infty.$$

因此 $\|f\|_\infty + \|g\|_\infty$ 是 $|f+g|$ 的一个界, 从而大于或等于最小的这种界即 $\|f+g\|_\infty$. 它给出了三角不等式. 其次我们注意到如果 $x \neq 0$, 则 b 为 $|f|$ 的界当且仅当 $|x|b$ 为 $|xf|$ 的界, 从而得到 $\|xf\|_\infty = |x|\|f\|_\infty$. 最后, $\|f\|_\infty \geqslant 0$ 且只有当 f 为零函数时才有 $\|f\|_\infty = 0$.

我们可以在上面讨论中用任意赋范线性空间 W 替换 $\mathbb{R}$. 函数 $f : A \to W$ 以 b 为界当且仅当对所有 A 中的 p 有 $\|f(p)\| \leqslant b$, 从而我们定义 $\mathfrak{B}(A, W)$ 上相应的一致范数为

$$\|f\|_\infty = \text{lub}\{\|f(p)\| : p \in A\}.$$

如果 $f \in \mathfrak{C}([0,1])$, 则我们知道连续函数 f 达到它值域的最小上界 (即 f"取其极大值"), 于是 $\|f\|_\infty$ 就是 $|f|$ 的极大值. 然而一般情形下, 必须用 lub 来给出定义.

球 记住了 $\|\alpha-\xi\|$ 被解释为 α 到 ξ 的距离则可自然地定义半径为 r, 关于中心 α 的开球为 $\{\xi : \|\alpha-\xi\| < r\}$. 我们记此球为 $B_r(\alpha)$. 平移 β 保持距离不变,

$$\|T_\beta(\xi) - T_\beta(\eta)\| = \|(\xi+\beta) - (\eta+\beta)\| = \|\xi - \eta\|,$$

因此 $\xi \in B_r(\alpha)$ 当且仅当 $\xi + \beta \in B_r(\alpha+\beta)$. 还有, 以 c 作纯量积则对所有的距离都乘以 c, 从而以相似的方式推出 $cB_r(\alpha) = B_{cr}(c\alpha)$.

尽管 $B_r(\alpha)$ 表现得像一个球, 但由不同范数定义的实际集合是各不相同的, 它们中有一些是 "看起来的非球状". 在图 3.2 中展示了三个范数 $\| \ \|_1, \| \ \|_2$, $\| \ \|_\infty$ 在 $\mathbb{R}^2$ 中以原点为中心的单位球.

一个 nls$V$ 的子集 A 是**有界的**是说, 如果它位于某个球中, 譬如 $B_r(\alpha)$ 中, 那么它也在一个以原点为中心的球中, 即 $B_{r+\|\alpha\|}(0)$ 中. 这只不过是如下的事实: 如果 $\|\xi-\alpha\| < r$ 则 $\|\xi\| < r + \|\alpha\|$; 它在重写 $\|\xi\|$ 为 $\|(\xi-\alpha)+\alpha\|$ 然后再运用三角不等式便能得到.

在向量 β 的与集合 A 不相接触的最大球的半径被自然地称作**从 β 到 A 的距离**. 很清楚它就是 $\mathrm{glb}\{\|\xi-\beta\| : \xi \in A\}$(见图 3.3).

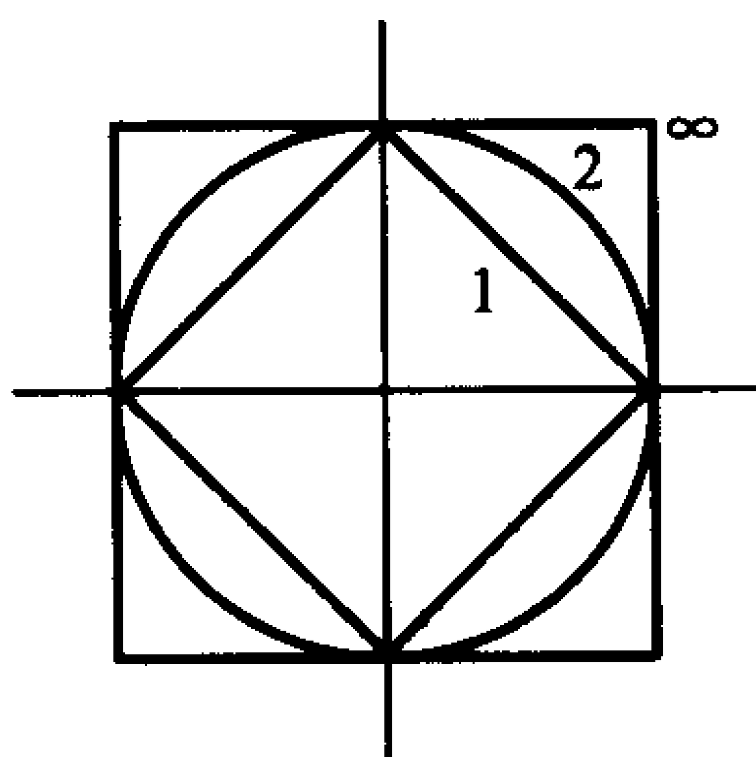


图 3.2

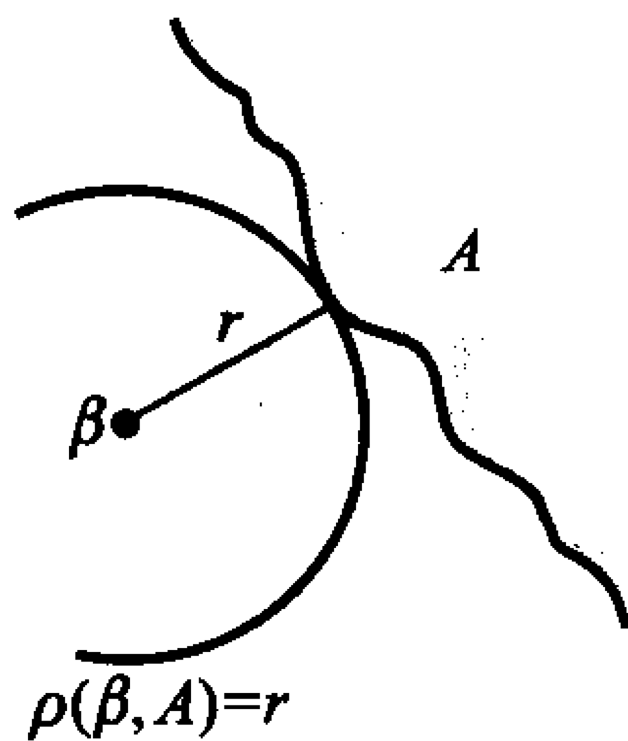


图 3.3

一个点 α 是集合 A 的**内点**是说在 α 有某个球被包含在 A 中. 这等价于说从 α 到 A 的补集的距离是正的 (但假定 A 不是整个 V), 它应该与读者的一个 "里面的" 点应该是什么的直观观念相一致. 称赋范线性空间的一个子集 A 是**开**的是说 A 的每个点均为内点.

如果我们的语言是前后一致的, 那么开球就应该是个开集. 的确如此: 如果 $\alpha \in B_r(\beta)$, 则 $\|\alpha-\beta\| < r$, 于是假如 $\delta \leqslant r - \|\alpha-\beta\|$ 并利用三角不等式, 我们便有 $B_\delta(\alpha) \subset B_r(\beta)$(见图 3.4). 读者应写出详细的证明. 他必须证明如果 $\xi \in B_\delta(\alpha)$, 则 $\xi \in B_r(\beta)$. 我们关于距离的直觉是十分值得信赖的, 但它们总是应该得到计算的检验. 通过动脑筋的推断读者大体能明白开集的任意集合的并集仍然是开

的. 特别, 任意一组开球的并仍是开的 (图 3.5), 这大体上是使开集形象化的最直观的方式.(见习题 2.9.)

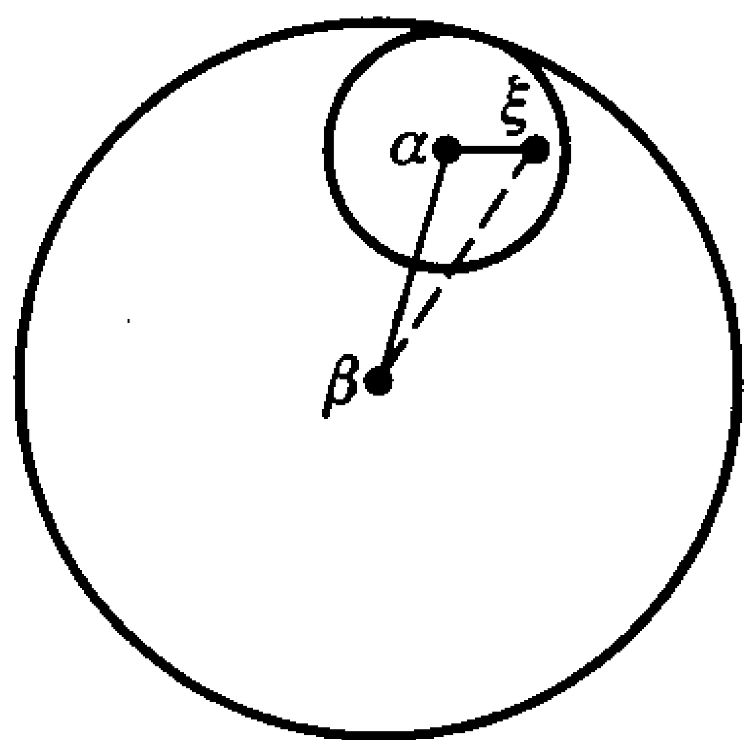
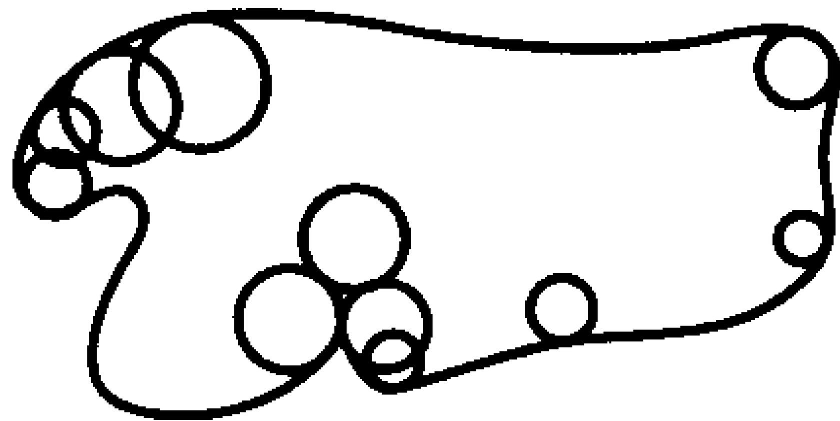


图 3.4


图 3.5

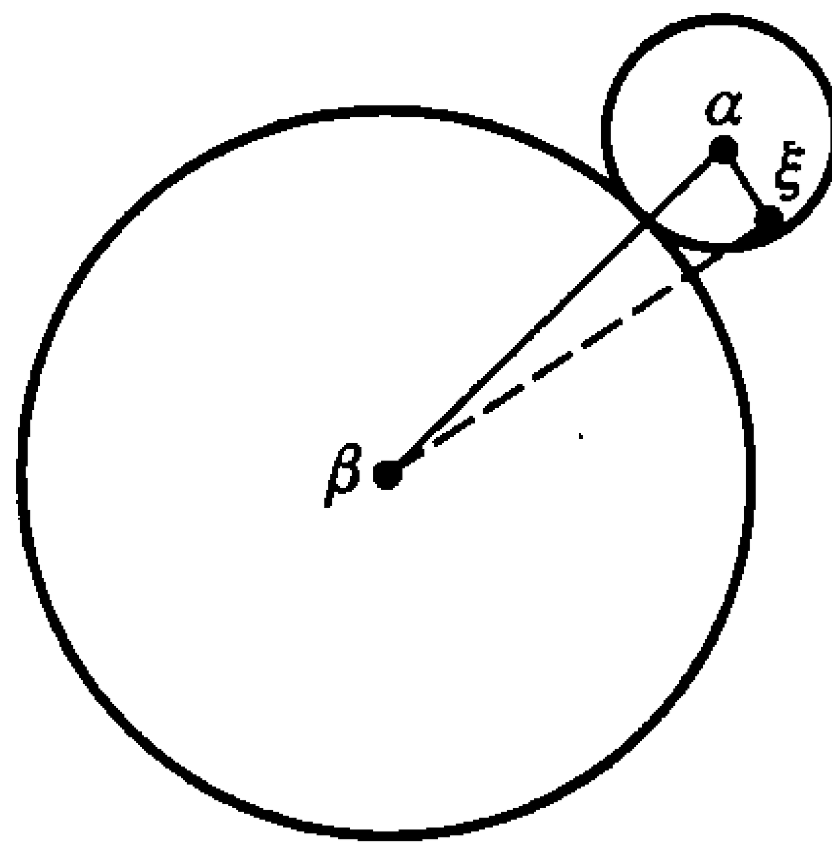


图 3.6

如果子集合 C 的补集 C' 是开的则称 C 为闭的.

我们上面的讨论表明一个非空集 C 是闭的当且仅当每个不在它里面的点与它的距离为正: $\alpha \notin C \Rightarrow \rho(\alpha, C) > 0$. 所谓在 β 的半径 r 的闭球 $B = \{\xi : \|\xi - \beta\| \leqslant r\}$ 是个闭集. 像图 3.6 所提示的那样, 其证明又是三角不等式的一个应用.

习题

2.1 证明如果 $\|\xi - \alpha\| \leqslant \|\alpha\|/2$, 则 $\|\xi\| \geqslant \|\alpha\|/2$.

2.2 详细证明

$$\|\boldsymbol{x}\|_1 = \sum_1^n |x_i|$$

是 $\mathbb{R}^n$ 上的一个范数. 还要证明

$$\|f\|_1 = \int_a^b |f(t)| dt$$

是 $\mathcal{C}([a, b])$ 上的一个范数.

2.3 对 $\mathbb{R}^n$ 中的 $\boldsymbol{x}$ 让 $|\boldsymbol{x}|$ 为欧氏长度

$$|x| = \left[\sum_1^n x_i^2\right]^{1/2},$$

另外设 (x, y) 为内积

$$(\boldsymbol{x}, \boldsymbol{y}) = \sum_1^n x_i y_i.$$

施瓦兹不等式说

$$|(\boldsymbol{x},\boldsymbol{y})| \leqslant |\boldsymbol{x}||\boldsymbol{y}|$$

且此不等式在 $\boldsymbol{x},\boldsymbol{y}$ 为相互无关时是严格的.

(a) 由平方再消去证明 $n=2$ 时的施瓦兹不等式.

(b) 现以相同方式对一般的 n 证明它.

2.4 继续上面的习题, 让明欧氏长度 $\boldsymbol{x}$ 是个范数. 关键的一步是三角不等式 $|\boldsymbol{x}+\boldsymbol{y}| \leqslant |\boldsymbol{x}|+|\boldsymbol{y}|$. 用平方和消去法把它化到施瓦兹不等式. 自然这就是我们的 2 - 范数 $\|\boldsymbol{x}\|_2$.

2.5 证明 $\mathbb{R}^2$ 中对范数 $\|\ \|_1$ 和 $\|\ \|_\infty$ 的单位球是图 3.2 所展示的那样.

2.6 证明开球是开集.

2.7 证明闭球是闭集.

2.8 给出 $\mathbb{R}^2$ 中一个既非开又非闭的一个子集的例子.

2.9 由开集的定义证明任何一个开集是一族 (可能非常大) 开球的并. 证明任一开集的并仍为开. 因此结论是一个集合为开当且仅当它是开球的并.

2.10 称一个赋范线性空间 V 的子集 A 是凸的是说, A 包含了它中任意两点的联结线段. 我们知道从 α 到 β 的线段是 $[0,1]$ 在映射 $t \mapsto t\beta+(1-t)\alpha$ 下的像. 因此 A 为凸当且仅当如果 $\alpha,\beta \in A$ 和 $t \in [0,1] \Rightarrow t\beta+(1-t)\alpha \in A$. 证明赋范线性空间中每个球 $B_r(\Gamma)$ 为凸.

2.11 一个半范数是除了把正性条件 n1 放宽为非负性:

$$\text{n1}'.\ p(\alpha) \geqslant 0 \text{ 对所有 } \alpha \text{ 成立}$$

外满足范数的其他所有条件. 因此 $p(\alpha)$ 可以对某些 α 为 0. 特别, 每个范数都是半范数. 证明:

(a) 如果 p 是向量空间 W 上的半范数, T 是从 V 到 W 的线性映射, 则 $p\circ T$ 是 V 上的半范数.

(b) $p\circ T$ 为范数当且仅当 T 为单射的且 p 为 T 的像上的范数.

2.12 证明两个半范数的和仍为半范数.

2.13 由上面的两个习题证明 (不是通过直接计算)

$$q(f)=\|f'\|_\infty+|f(t_0)|$$

是 $[a,b]$ 上所有连续可微实函数的空间 $\mathfrak{C}^1([a,b])$ 上的半范数, 其中 t_0 是 $[a,b]$ 中一个固定点. 证明 q 是一个范数.

2.14 证明两个有界集的和是有界的.

2.15 证明和 $B_r(\alpha)+B_s(\beta)$恰好是球 $B_{r+s}(\alpha+\beta)$.

3.3 连续性

设 V 和 W 为任意两个赋范线性空间. 我们对两个范数都用 $\|\ \|$ 表示. 这种含糊的用法不会引起混乱. 它像是用 "0" 表示所考虑的全部向量空间的零元素那种样子的含糊. 如果在早先我们给出单实变量的实函数极限的定义中用一般的范数符号 $\|\ \|$ 替换绝对值记号 $|\ |$, 则它逐字逐句地成为在一般架构中相应的收敛性定义. 但是我们仍将重述此定义并借此机会放宽对 f 定义域的假定条件. 因此, 设 A 为 V 的任一子集, 并设 f 为从 A 到 W 的任一映射.

定义 3.1 *我们说当 ξ 趋向 α 时 $f(\xi)$ 趋向 β, 并记为当 $\xi \to \alpha$ 时 $f(\xi) \to \beta$ 指的是, 如果对每个 ε 存在一个 δ 使得*

$$\xi \in A \ \text{和}\ 0 < \|\xi - \alpha\| < \delta \Rightarrow \|f(\xi) - \beta\| < \varepsilon.$$

如果 $\alpha \in A$, 并且当 $\xi \to \alpha$ 时 $f(\xi) \to f(\alpha)$, 我们则说 f 在 α 是*连续的*. 于是我们可以去掉 $\xi \neq \alpha$ 的要求, 并有了连续性的直接的 ε, δ - 特征描述: 如果对每个 ε 存在一个 δ 使得 $\|\xi - \alpha\| < \delta \Rightarrow \|f(\xi) - f(\alpha)\| < \varepsilon$, 则说 f 在 α 是连续的. 在这里可以理解为 ξ 是在 f 的定义域 A 上的全词量化. 如果 f 在其定义域中每点 α 均连续则说 f 是*连续的*. 如果把一个数的绝对值换作一个向量的范数, 则在第 1 节中作为实例的极限定理对于赋范线性空间逐字逐句地成立. 我们将要求读者写出习题中这些翻录文字中的几个.

有一个比在 α 的连续性更强的性质, 如果它成立则在使用时要简单得多. f 在 α 是*利普希茨连续的*是说, 如果存在常数 c 使得 $\|f(\xi) - f(\alpha)\| \leqslant c\|\xi - \alpha\|$ 对所有充分靠近 α 的 ξ 成立. 那就是说, 存在常数 c 和 r 使得

$$\|\xi - \alpha\| < r \Rightarrow \|f(\xi) - f(\alpha)\| \leqslant c\|\xi - \alpha\|.$$

重要之处在于我们现在可以简单地取 δ 为 ε/c 了 (假定 ε 足够小使得 $\delta \leqslant r$; 不然我们不得不令 $\delta = \min\{\varepsilon/c, r\}$). 如果在 A 对所有 ξ, η 存在一个常数 c 使得 $\|f(\xi) - f(\eta)\| \leqslant c\|\xi - \eta\|$, 我们则说 f 是一个*利普希茨函数*(在它的定义域 A 上). 对于*线性*映射 $T: V \to W$, 利普希茨不等式被更为简洁地写为

$$\|T(\zeta)\| \leqslant c\|\zeta\|,$$

对所有 $\zeta \in V$ 成立; 我们刚才使用的事实现在是 $T(\xi) - T(\eta) = T(\xi - \eta)$, 从而可令 $\zeta = \xi - \eta$. 在这方面我们按习惯称 T 为*有界*线性映射而不是利普希茨线性映射, 而且称任意这样的 c 为 T 的一个*界*.

由初等微积分我们知道, 如果 f 是 $[a,b]$ 上的连续实函数 (即如果 $f \in \mathfrak{C}([a,b])$), 则 $|\int_a^b f(x)dx| \leqslant m(b-a)$, 其中 m 是 $|f(x)|$ 的极大值. 但这正好是 f 的一致范数, 故此不等式可重写为 $|\int_a^b f| \leqslant (b-a)\|f\|_\infty$. 这表明如果在 $\mathfrak{C}([a,b])$ 上使用一致范数, 则 $f \to \int_a^b f$ 是一个有界线性泛函, 其界为 $b-a$.

应该立即指出这与我们早先讨论过的有界性的概念不是同一回事. 在那里我们称一个实函数有界是说它的值域是 $\mathbb{R}$ 中的一个有界子集合. 在这里的类比应当被称作有界的向量值函数, 如果它的值域是范数有界的. 然而一个非零的线性变换不可能在这种意义下有界, 这是因为

$$\|T(x\alpha)\| = |x|\|T(\alpha)\|.$$

现在的定义等于商$T(\alpha)/\|\alpha\|$(在 $V - \{0\}$ 上) 在原先意义下的有界性. 对一个线性映射, 原来连续和利普希茨是同一回事.

定理 3.1 *设 T 为从赋范线性空间 V 到赋范线性空间 W 的一个线性映射. 则下列条件等价:*

(1) T 在一点连续;

(2) T 连续;

(3) T 有界.

证明 (1) $\Rightarrow$ (3). 设 T 在 α_0 连续. 取 $\varepsilon = 1$, 于是存在 δ 使得 $\|\alpha - \alpha_0\| < \delta \Rightarrow \|T(\alpha) - T(\alpha_0)\| < 1$. 令 $\xi = \alpha - \alpha_0$, 并运用 T 的可加性, 我们有 $\|\xi\| < \delta \Rightarrow \|T(\xi)\| < 1$. 现对任意非零 η, 让 $\xi = \delta\eta/2\|\eta\|$, 它的范数为 $\delta/2$. 因此, $\|T(\xi)\| < 1$. 但 $\|T(\xi)\| = \delta\|T(\eta)\|/2\|\eta\|$, 它给出了 $\|T(\eta)\| < 2\|\eta\|/\delta$. 因此 T 以 $C = 2/\delta$ 为界.

(3) $\Rightarrow$ (2). 设 $\|T(\xi)\| \leqslant C\|\xi\|$ 对所有 ξ 成立. 于是对任意 α_0 和任意 ε 我们可取 $\delta = \varepsilon/C$, 从而有

$$\|\alpha - \alpha_0\| < \delta \Rightarrow \|T(\alpha) - T(\alpha_0)\| = \|T(\alpha - \alpha_0)\| \leqslant C\|\alpha - \alpha_0\| < C\delta = \varepsilon.$$

(2) $\Rightarrow$ (1). 平凡的. □

在下面的引理中我们将证明范数函数是 V 到 $\mathbb{R}$ 的一个利普希茨函数.

引理 3.1 *对所有 $\alpha, \beta \in V$, $|\,\|\alpha\| - \|\beta\|\,| \leqslant \|\alpha - \beta\|$.*

证明 我们有 $\|\alpha\| = \|(\alpha - \beta) + \beta\| \leqslant \|\alpha - \beta\| + \|\beta\|$, 故 $\|\alpha\| - \|\beta\| \leqslant \|\alpha - \beta\|$. 相似地, 有 $\|\beta\| - \|\alpha\| \leqslant \|\beta - \alpha\| = \|\alpha - \beta\|$. 这一对不等式等价于此引理. □

当我们研讨具有连续微分的映射时还将出现其他的利普希茨映射. 粗略地说, 利普希茨性质处于连续性和连续的可微性之间, 它常常是我们在连续可微性的假设下实际应用的条件.

一个有界线性变换 T 的最小界被称作它的范数. 即

$$\|T\| = \text{lub}\{\|T(\alpha)\|/\|\alpha\| : \alpha \neq 0\}.$$

例如, 设 $T : \mathfrak{C}([a,b]) \to \mathbb{R}$ 为黎曼积分, $T(f) = \int_a^b f(x)dx$. 我们在前面见过, 如果在 $\mathfrak{C}([a,b])$ 上采用一致范数 $\|f\|_\infty$ 则 T 以 $b-a$ 为界:$|T(f)| \leqslant (b-a)\|f\|_\infty$. 另一方面, 不存在更小的界, 这是因为 $\int_a^b 1 = (b-a) = (b-a)\|1\|_\infty$. 因此 $\|T\| = b-a$. 上述定义有另一种有用的公式化的阐述. 因为由齐性

$$\|T(\alpha)\|/\|\alpha\| = \|T(\alpha/\|\alpha\|)\|,$$

并且 $\beta = \alpha/\|\alpha\|$ 具有范数 1, 从而我们有

$$\|T\| = \text{lub}\{\|T(\beta)\| : \|\beta\| = 1\}.$$

最后, 如果 $\|\gamma\| \leqslant 1$, 则 $\gamma = x\beta$, 其中 $\|\beta\| = 1, |x| \leqslant 1$, 且

$$\|T(\gamma)\| = |x|\|T(\beta)\| \leqslant \|T(\beta)\|.$$

因此我们有了一个效果较差但仍是有用的特征描述:

$$\|T\| = \text{lub}\{\|T(\gamma)\| : \|\gamma\| \leqslant 1\}.$$

最后的这两个公式化陈述的是*一致范数*. 因此, 如果 B_1 是单位闭球 $\{\xi : \|\xi\| \leqslant 1\}$, 我们看出一个线性的 T 为有界当且仅当 $T \upharpoonright B_1$ 在老的意义下有界, 于是

$$\|T\| = \|T \upharpoonright B_1\|_\infty.$$

一个线性映射 $T : V \to W$ 如果对 V 中所有 ξ 有 $\|T(\xi)\| \geqslant b\|\xi\|$ 则是*被 b 界于下的*.如果 T 有一个有界的逆, 且 $m = \|T^{-1}\|$, 则 T 被 $1/m$ 界于下, 这是因为对所有 $\eta \in W$, $\|T^{-1}(\eta)\| \leqslant m\|\eta\|$ 当且仅当 $\|\xi\| \leqslant m\|T(\xi)\|$ 对所有 $\xi \in V$ 成立.

如果 V 是有限维的, 这是对的; 反之, 如果 T 是界于下的则它是可逆的 (为什么?), 但一般情形不对.

如果 V 和 W 为赋范线性空间, 则 $\text{Hom}(V, W)$ 被定义为所有*有界*线性映射 $T : V \to W$ 的集合. 2.3 节中所有结果仍然正确, 然而我们还需要另外一些论断.

定理 3.2 如果 $\|T\|$ 按上面那样定义为 T 的最小界, 则 $\mathrm{Hom}(V,W)$ 自身是一个赋范线性空间.

证明 利用等式 $\|T\| = \|T \upharpoonright B_1\|_\infty$ 由第 2 节关于一致范数的讨论推出. □

定理 3.3 如果 U,V 和 W 都是赋范线性空间, 如果 $T \in \mathrm{Hom}(U,V)$, $S \in \mathrm{Hom}(V,W)$, 则 $S \circ T \in \mathrm{Hom}(U,W)$ 且 $\|S \circ T\| \leqslant \|S\|\|T\|$. 从而在右边复合一个固定的 T 是从 $\mathrm{Hom}(V,W)$ 到 $\mathrm{Hom}(U,W)$ 的一个有界线性变换, 对于从左边复合一个固定的 S 也有相似的结果.

证明

$$\|(S \circ T)(\alpha)\| = \|S(T(\alpha))\| \leqslant \|S\|\|T(\alpha)\| \leqslant \|S\|(\|T\| \cdot \|\alpha\|) = (\|S\| \cdot \|T\|)(\|\alpha\|).$$

因此 $S \circ T$ 以 $\|S\| \cdot \|T\|$ 为界, 其余的一切也立即被推导出来. □

像从前那样, 共轭空间 V^* 为 $\mathrm{Hom}(V,\mathbb{R})$, 它现在是所有有界线性泛函的空间.

习题

3.1 写出下列极限定理的 ε,δ - 证明.

(1) 设 V 和 W 为赋范线性空间, F 和 G 为 V 到 W 的映射. 如果 $\lim_{\xi\to\alpha}F(\xi) = \mu, \lim_{\xi\to\alpha}G(\xi) = \nu$, 则 $\lim_{\xi\to\alpha}(F+G)(\xi) = \mu+\nu$.

(2) 已知 $F: V \to W$, $g: V \to \mathbb{R}$, 如果当 $\xi \to \alpha$ 时 $F(\xi) \to \mu$, $g(\xi) \to b$, 则 $(gF)(\xi) \to b\mu$.

3.2 证明当 $\xi \to \alpha$ 时如果 $F(\xi) \to \mu$, $\eta \to \mu$ 时 $G(\eta) \to \lambda$, 则当 $\xi \to \alpha$ 时 $G \circ F(\xi) \to \lambda$. 给出你所证明的这个定理的仔细而完整的陈述.

3.3 假设 A 是一个 nls$V$ 的开子集, $\alpha_0 \in A$. 假设 $F: A \to \mathbb{R}$ 使得 $\lim_{\alpha\to\alpha_0}F(\alpha) = b \neq 0$. 证明当 $\alpha \to \alpha_0$ 时 $1/F(\alpha) \to 1/b$(用 ε,δ - 证明).

3.4 函数 $f(x) = |x|^r$ 对任意正数 r 在 $x = 0$ 为连续. 证明如果 $r < 1, f$ 在 $x = 0$ 不是利普希茨连续的. 但是, 证明在 $x = a$, 如果 $a > 0$ 时 f 是利普希茨连续的 (用中值定理).

3.5 用微积分的中值定理和导数的定义证明如果 f 是某区间 I 上的实函数且 f' 处处存在, 则 f 为利普希茨映射当且仅当 f' 为有界函数. 证明 $\|f'\|_\infty$ 是一最小的利普希茨常数 C.

3.6 $\|T\|$ 的 "工作规则" 是

(1) $\|T(\xi)\| \leqslant \|T\|\|\xi\|$ 对所有 ξ 成立;

(2) $\|T(\xi)\| \leqslant b\|\xi\|$, 对所有 $\xi \Rightarrow \|T\| < b$.
证明这些规则.

3.7 证明如果我们在 $\mathbb{R}^n$ 上用 1 - 范数 $\|\boldsymbol{x}\|_1 = \sum_1^n |x_i|$, 则线性泛函

$$L_{\boldsymbol{a}} : \boldsymbol{x} \to \sum_1^n a_i x_i$$

的范数为 $\|\boldsymbol{a}\|_\infty$.

3.8 相似地证明如果 $\|\boldsymbol{x}\| = \|\boldsymbol{x}\|_\infty$ 则 $\|L_{\boldsymbol{a}}\| = \|\boldsymbol{a}\|_1$.

3.9 用上面习题证明, 如果在 $\mathbb{R}^n$ 上 $\|\boldsymbol{x}\|$ 是 1 - 范数, 则

$$\boldsymbol{x} = \text{lub}\{|f(x)| : f \in (\mathbb{R}^n)^* \text{且} \|f\| \leqslant 1\}.$$

3.10 证明如果 T 在 $\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 中有矩阵 $\boldsymbol{t} = \{t_{ij}\}$, 且如果我在 $\mathbb{R}^n$ 中采用 1 - 范数 $\|\boldsymbol{x}\|_1$ 而在 $\mathbb{R}^m$ 中用一致范数 $\|\boldsymbol{y}\|_\infty$, 则 $\|T\| = \|\boldsymbol{t}\|_\infty$.

3.11 给出一个不是有界的线性映射的例子以证明 'Hom(V, W)' 的意义已经改变. 在正文中有一个这种例子.

3.12 对 V 中一个固定的 ξ 定义映射 $ev_\xi : \text{Hom}(V, W) \to W$ 为 $ev_\xi(T) = T(\xi)$. 证明 ev_ξ 是个有界线性映射.

3.13 在上面的习题中实际上成立 $\|ev_\xi\| = \|\xi\|$, 但要证明这点我们需要一个新定理:

定理 3.4 给出赋范线性空间 V 中的 ξ, 则存在 V^* 中一个泛函 f 使得 $\|f\| = 1$, 且 $|f(\xi)| = \|\xi\|$. 假定此定理成立, 证明 $\|ev_\xi\| = \|\xi\|$.[提示: 设若你已经证明了 $\|ev_\xi\| \leqslant \|\xi\|$. 现在你需要一个 $\text{Hom}(V, W)$ 中的 T 使得 $\|T\| = 1$ 而 $\|T(\xi)\| = \|\xi\|$. 考虑一个合适的并向量.]

3.14 设 $\boldsymbol{t} = \{t_{ij}\}$ 是个方阵并定义 $\|\boldsymbol{t}\|$ 为 $\max_i(\sum_j |t_{ij}|)$. 证明这是所有 $n \times n$ 矩阵的空间 $\mathbb{R}^{\overline{n} \times \overline{n}}$ 上的一个范数. 证明 $\|\boldsymbol{st}\| \leqslant \|\boldsymbol{s}\| \cdot \|\boldsymbol{t}\|$. 计算单位矩阵的范数.

3.15 设 V 是 $\mathbb{R}^n$ 在一致范数 $\|\boldsymbol{x}\|_\infty = \max\{|x_i|\}$ 下的赋范线性空间. 如果 $T \in \text{Hom}V$, 证明在上个习题中定义 $\|\boldsymbol{t}\|$ 是 $\|T\|$ 的范数, 其中 $\boldsymbol{t}$ 是 T 的矩阵. 那就是证明

$$\|T\| = \max_i \left[\sum_{j=1}^n |t_{ij}|\right].$$

(首先证明 $\|\boldsymbol{t}\|$ 是 T 的一个上界, 然后证明 $\|T(\boldsymbol{x})\| = \|\boldsymbol{t}\|\|\boldsymbol{x}\|$, 其中 $\boldsymbol{x}$ 是一个特别选定的向量.) 是否上面的习题中有些部分现在显得多余了?

3.16 假设有下面的事实: 如果 $f \in \mathfrak{C}([0,1])$ 及 $\|f\|_1 = a$, 则给出 ε, 存在一个函数 $g \in \mathfrak{C}([0,1])$ 使得

$$\|g\|_\infty = 1 \quad 且 \quad \int_0^1 fg > a - \varepsilon.$$

设 $K(s,t)$ 在 $[0,1]\times[0,1]$ 上连续并以 b 为界. 定义 $T : \mathfrak{C}([0,1]) \to \mathfrak{B}([0,1])$ 为 $Th = k$, 其中

$$k(s) = \int_0^1 K(s,t)h(t)dt.$$

如果 V 和 W 分别是在一致范数下的赋范线性空间 $\mathfrak{C}$ 和 $\mathfrak{B}$, 证明

$$\|T\| = \operatorname*{lub}_s \int |K(s,t)|dt.$$

(提示: 如同上一个习题那样进行.)

3.17 设 V 和 W 为赋范线性空间, A 为 V 中含有多于一个点的任一子集. 让 $\mathfrak{L}(A,W)$ 为 A 到 W 的所有利普希茨映射的集合. 对 $\mathfrak{L}(A,W)$ 中的 f, 让 $p(f)$ 为 f 的最小利普希茨常数. 那就是说,

$$p(f) = \operatorname*{lub}_{\xi\neq\eta} \frac{\|f(\xi) - f(\eta)\|}{\|\xi - \eta\|}.$$

证明 $\mathfrak{L}(A,W)$ 是个向量空间 V 而 p 是 V 上的半范数.

3.18 继续上面的习题, 证明如果 α 是 A 中任一固定点, 则 $p(f) + \|f(\alpha)\|$ 是 V 上的范数.

3.19 设 K 是从一个赋范线性空间 V 的一个子集 A 到 V 的一个映射, 它与恒同映射相差一个具有小于 1 的常数 c 的利普希茨映射. 我们不妨取 $c = \frac{1}{2}$, 于是我们的假说便是

$$\|K(\xi) - K(\eta) - (\xi - \eta)\| \leqslant \frac{1}{2}\|\xi - \eta\|.$$

证明 K 是单射的而其逆是一个具有常数 2 的利普希茨映射.

3.20 继续上面的习题, 假设另外有 K 的定义域 A 为 V 的一个开子集, 并且对 A 中任一闭球 C, $K[C]$ 是个闭集. 证明如果半径 r 关于 α 的闭球 $C = C_r(\alpha)$ 是 A 中的一个子集, 则 $K[C]$ 包含了球 $B = B_{r/7}(\gamma)$, 其中 $\gamma = K(\alpha)$. 这个证明虽然初等却富于技巧. 如果有 B 中点 v 不在 $K[C]$ 中, 则由于 $K[C]$ 为闭集, 有一个关于 v 的与 $K[C]$ 不相交的最大的球 B', 以及 $K[C]$ 中一个点 $\eta = K(\xi)$ 可以按我们的要求靠近 B'. 现在如果我们加上 $v - \eta$ 来改变 ξ, 则 K 的值的变化将逼近 $v - \eta$ 充分地靠近, 使得 K 的新值在 B' 中. 如果我们也能够证明新的值 $\xi + (v - \eta)$ 在 C 中, 则 K 的新值便在 $K[C]$ 中了, 从而得到矛盾.

画一幅图. 显然 B' 的半径 ρ 最多为 $r/7$. 证明如果选取 $\eta = K(\xi)$ 使得 $\|v - \eta\| \leqslant 3/2\rho$, 则上面的论断来自三角不等式和在习题 3.19 中所显示的利普希茨不等式. 你必须要证明

$$\|K(\xi + (v - \eta)) - v\| < \rho$$

和

$$\|(\xi + (v - \eta)) - \alpha\| < r.$$

3.21 假定上面习题的结果成立, 证明

$$B_{r/8}(\gamma) \subset K[B_r(\alpha)].$$

因此而证明 $K[A]$ 是 V 中的一个开子集. 叙述一个关于 K 的利普希茨可逆性的定理, 包括用到的在上面习题中的对 K 的所有假设条件.

3.22 在下一章我们将看到如果 V 和 N 是有限维空间, 则任何从 V 到 W 的连续映射把有界闭集带到有界闭集. 承认这一点和上面习题的结果, 证明下面的定理.

定理 3.5 *设 F 是由一个有限维赋范线性空间 V 的一个开子集 A 到有限维赋范线性空间 W 的一个映射. 假设在 $\mathrm{Hom}(V, W)$ 中存在一个 T 使得 T^{-1} 存在, 并且 $F - T$ 在 A 上为利普希茨, 常数为 $1/2m$, 其中 $m = \|T^{-1}\|$. 则 F 为单射, 其值域 $R = F[A]$ 是 W 中的开子集并且逆 F^{-1} 是利普希茨连续的, 且有常数 $2m$.*

3.4 等价的范数

两个赋范线性空间 V 和 W 为范数同构的是说如果有一个 V 到 W 的双射 T 使得 $T \in \mathrm{Hom}(V, W)$ 及 $T^{-1} \in \mathrm{Hom}(W, V)$. 就是说, 一个同构是一个线性的同构 T 使得 T 和 T^{-1} 都是连续的 (有界的). 像通常那样, 我们看待同构的空间在本质上是相同的. 对同一空间上两个不同的范数我们有下面的定义.

定义 4.1 *在同一向量空间 V 上两个范数 p 和 q 被称为等价的是说, 如果存在常数 a 和 b 使得 $p \leqslant aq$ 和 $q \leqslant bp$.*

于是 $(1/b)q \leqslant p \leqslant aq$ 和 $(1/a)p \leqslant q \leqslant bp$, 故两个范数为等价当且仅当两个中任一个被另一个的两个倍数包围. 上面的定义只不过是说从 V 到 V 的恒等映射 $\xi \to \xi$ 当被考虑为赋范线性空间 $\prec V, p \succ$ 到赋范空间 $\prec V, q \succ$ 的映射时在两个方向上都是有界的, 因而这两个赋范线性空间是同构的.

如果 V 是无穷维的, 两个范数一般不是等价的. 例如, 如果 $V = \mathfrak{C}([0,1])$, $f_n(t) = t^n$, 则 $\|f_n\|_1 = 1/n+1$, 而 $\|f_n\|_\infty = 1$. 因而没有常数 a 使得 $\|f\|_\infty \leqslant a\|f\|_1$ 对所有 $f \in \mathfrak{C}([0,1])$ 成立, 从而范数 $\|\ \|_\infty$ 和 $\|\ \|_1$ 在 $V = \mathfrak{C}([a,b])$ 上不等价. 这就是为什么一个赋范线性空间的概念依赖于对所给范数的假定.

然而我们有下面的定理, 我们将在下一章中用比现在运用的更精妙的方法加以证明.

定理 4.1 在一个有限维向量空间 V 上, 所有的范数都是等价的.

在本章中我们将需要这个定理以及偶尔需要下面它的推论.

定理 4.2 如果 V 和 W 是有限维赋范线性空间, 于是从 V 到 W 的每个线性映射 T 必是有界的.

证明 因为上面的定理, 只要证明 T 对于某一对范数是有界的就够了. 设 $\theta:\mathbb{R}^n\to V$ 和 $\varphi:\mathbb{R}^n\to W$ 为任意的基同构, 并让 $\{t_{ij}\}$ 为 $\overline{T}=\varphi^{-1}\circ T\circ\theta\in\mathrm{Hom}(\mathbb{R}^n,\mathbb{R}^m)$ 的矩阵. 于是

$$\begin{aligned}\|\overline{T}\boldsymbol{x}\|_\infty&=\max_i\left|\sum_j t_{ij}x_j\right|\leqslant\left(\max_{i,j}|t_{ij}|\right)\left(\sum_1^n|x_j|\right)\\&=b\|\boldsymbol{x}\|_1,\end{aligned}$$

其中 $b=\max|t_{ij}|$. 现令 $q(\eta)=\|\varphi^{-1}(\eta)\|_\infty$ 和 $p(\xi)=\|\theta^{-1}(\xi)\|_1$ 各为 W 和 V 上的范数, 这由引理 2.1 得知; 又因为

$$q(T(\xi))=\|\overline{T}(\theta^{-1}\xi)\|_\infty\leqslant b\|\theta^{-1}\xi\|_1=bp(\xi),$$

我们看出 T 在 V 和 W 上以 p,q 为范数时以 b 为界. □

如果我们改变为一个等价的范数, 我们则仅仅经过了一个同构, 从而所有连续的线性性质并没有改变. 例如:

定理 4.3 向量空间 $\mathrm{Hom}(V,W)$ 在或者定义域的范数或者值域的范数换作一个等价范数时仍然是一样的, 并这两个在 $\mathrm{Hom}(V,W)$ 诱导出的范数是等价的.

证明 留给读者. □

我们现在要问, 在两个赋范线性空间的笛卡儿积 $V\times W$ 上我们要给出什么类型的范数。自然的是想选取乘积的范数为使得与积空间的两个因子空间相关的那些基本映射应该是连续的; 这里所说的基本映射是两个投影 π_i 和两个内(单) 射 θ_i. 结果表明这些要求在等价范围内惟一决定了这个乘积范数. 因为如果 $\|\prec\alpha,\xi\succ\|$ 具有这些性质, 则

$$\begin{aligned}\|\prec\alpha,\xi\succ\|&=\|\prec\alpha,0\succ+\prec0,\xi\succ\|\leqslant\|\prec\alpha,0\succ\|+\|\prec0,\xi\succ\|\\&\leqslant k_1\|\alpha\|+k_2\|\xi\|\leqslant k(\|\alpha\|+\|\xi\|),\end{aligned}$$

其中 k_i 为内射 θ_i 的一个界而 k 是 k_1 和 k_2 中较大者. 另外, $\|\alpha\| \leqslant c_1 \| \prec \alpha, \xi \succ \|, \|\xi\| \leqslant c_2 \| \prec \alpha, \xi \succ \|$, 这是由投影 π_i 的有界性决定的; 故 $\|\alpha\| + \|\xi\| \leqslant c \| \prec \alpha, \xi \succ \|$, 其中 $c = c_1 + c_2$. 现在清楚看出 $\|\alpha\| + \|\xi\|$ 是 $V \times W$ 上一个范数 $\| \|_1$, 而我们前面的论证表明 $\| \prec \alpha, \xi \succ \|$ 满足我们的要求当且仅当它等价于 $\| \|_1$. 任何一个这样的范数都称作 $V \times W$ 的一个**乘积范数**. 最常用的乘积范数是一致 (乘积) 范数

$$\| \prec \alpha, \xi \succ \|_\infty = \max\{\|\alpha\|, \|\xi\|\},$$

欧几里得 (乘积) 范数 $\| \prec \alpha, \xi \succ \|_2 = (\|\alpha\|^2 + \|\xi\|^2)^{1/2}$ 和上面的和 (乘积) 范数 $\| \prec \alpha, \xi \succ \|_1$. 我们把验证一致和欧氏范数确为范数作为习题留给读者.

这三个乘积范数中每一个都能对 n 个因子的空间定义, 情况像两个因子那样, 我们把对这个一般情形的事实集中为一个定理.

定理 4.4 如果 $\{\prec V_i, p_i \succ\}_1^n$ 为赋范线性空间的有限集合, 则 $\| \|_1, \| \|_2$ 和 $\| \|_\infty$ 在 $V = \prod_{i=1}^n V_i$ 被定义为 $\|\boldsymbol{\alpha}\|_1 = \sum_1^n p_i(\alpha_i), \|\boldsymbol{\alpha}\|_2 = (\sum_1^n p_i(\alpha_i)^2)^{1/2}$, 以及 $\|\boldsymbol{\alpha}\|_\infty = \max\{p_i(\alpha_i) : i = 1, \cdots, n\}$ 为 V 上的等价范数, 而且其中每个均是乘积范数, 其意义是说所有的投影 π_i 和内射 θ_i 均为连续.

* 上面看起来好像我们在做的一切就是在 $\mathbb{R}^n$ 上取任意一个范数 $\| \|$, 然后在乘积空间 V 上定义一个范数 $||| \; |||$ 为

$$|||\boldsymbol{\alpha}||| = \| \prec p_1(\alpha_1), \cdots, p_n(\alpha_n) \succ \|.$$

这差不多是对的. 有兴趣的读者会发现, 在 $\mathbb{R}^n$ 上的 $\| \|$ 必须有这样的性质, 即如果 $|x_i| \leqslant |y_i|, i = 1, \cdots, n$, 则 $\|\boldsymbol{x}\| \leqslant \|\boldsymbol{y}\|$, 这是导出 V 中 $||| \; |||$ 的三角不等式所需要的. 如果我们称 $\mathbb{R}^n$ 上这样的范数为**递增范数**, 则下面的陈述成立.

如果 $\| \|$ 为 $\mathbb{R}^n$ 上的任何一个递增范数, 则 $|||\boldsymbol{\alpha}||| = \| \prec p_1(\alpha_1), \cdots, p_n(\alpha_n) \succ \|$ 为 $V = \prod_1^n V_i$ 上的一个乘积范数.

但是在本书中我们将只使用 1 -,2 -,∞ - 乘积范数.*

三角不等式, 加法的连续性以及我们对乘积范数的要求形成了一组近乎等价的条件. 特别地, 我们进行下面的观察.

引理 4.1 如果 V 为赋值线性空间, 则加法运算是从 $V \times V$ 到 V 的有界线性映射.

证明 V 上范数的三角不等式恰好是说加法以 1 为界, 其中在 $V\times V$ 上用了和范数. □

如果映射 $\prec x_1,\cdots,x_n\succ\mapsto\sum_1^n x_i$ 是个从 $\prod_1^n V_i$ 到 V 的范数同构则说赋范线性空间 V 是一个 (范数) 直和 $\bigoplus_1^n V_i$. 这就是说, V 上所给出的范数必定等价于当将它看做 $\prod_1^n V_i$ 时所获得的乘积范数. 如 V 代数地等于 $\bigoplus_1^n V_i$, 我们由 V 上范数的三角不等式总有

$$\|x\|=\|\sum_1^n x_i\|\leqslant\sum_1^n\|x_i\|,$$

其中右端的和是 $\prod_1^n V_i$ 的 1 - 范数. 因此如果相反地存在一个 n 联组的常数 $\{k_i\}$ 使得对所有 x 有 $\|x_i\|\leqslant k_i\|x\|$ 则 V 将是范数直和 $\bigoplus_1^n V_i$. 这等于说所有射影 $P_i:x\mapsto x_i$ 为有界. 因此

定理 4.5 *如果 V 为赋范线性空间, 且 V 代数地等于直和 $V=\bigoplus_1^n V_i$, 则作为赋范线性空间 $V=\bigoplus_1^n V_i$ 当且仅当其相伴的射影 $\{P_i\}$ 全都有界.*

习题

4.1 在范数以等价的范数替换时 $\mathrm{Hom}(V,W)$ 不变这个事实可以看作定理 3.3 的推论. 证明其确实如此.

4.2 写出一串十分显见的不等式以表明 $\mathbb{R}^n$ 上的 $\|\ \|_1$, $\|\ \|_2$ 和 $\|\ \|_\infty$ 等价. 讨论当 $n\to\infty$ 时发生的情况.

4.3 设 V 为 n 维向量空间, 并考虑 V 上所有形如 $p\circ\theta$ 的范数汇集, 其中 $\theta:V\to\mathbb{R}^n$ 为坐标同构, p 为 $\mathbb{R}^n$ 上范数 $\|\ \|_1$, $\|\ \|_2$, $\|\ \|_\infty$ 中的一个. 证明这些范数全都是等价的. (利用上面的习题以及定理 4.2 的推理.)

4.4 证明 $\|\prec\alpha,\xi\succ\|=\max\{\|\alpha\|,\|\xi\|\}$ 是 $V\times W$ 上的范数.

4.5 证明 $\|\prec\alpha,\xi\succ\|=\|\alpha\|+\|\xi\|$ 是 $V\times W$ 上的范数.

4.6 证明 $\|\prec\alpha,\xi\succ\|=(\|\alpha\|^2+\|\xi\|^2)^{1/2}$ 是 $V\times W$ 上的范数.

4.7 假定有习题 4.4 到 4.6, 用归纳法证明定理 4.4 的相应部分.

4.8 证明如果 A 是 $V\times W$ 中一个开子集则 $\pi_1(A)$ 是 V 的一个开子集.

4.9 以 (ε,δ) 证明 $\prec T,S\succ\to S\circ T$ 是从

$$\mathrm{Hom}(V_1,V_2)\times\mathrm{Hom}(V_2,V_3)\quad\text{到}\quad\mathrm{Hom}(V_1,V_3)$$

的一个连续映射, 其中每个 V_i 都是赋范线性空间.

4.10 设 $\|\ \|$ 为 $\mathbb{R}^n$ 上任一个递增范数; 即如果对所有 $i, x_i \leqslant y_i$, 则 $\|\boldsymbol{x}\| \leqslant \|\boldsymbol{y}\|$. 设 p_i 为向量空间 V_i 上的范数, $i = 1, \cdots, n$. 证明

$$|||\boldsymbol{\alpha}||| = \| \prec p_1(\alpha_1), \cdots, p_n(\alpha_n) \succ \|$$

是 $V = \prod_1^n V_i$ 上的一个范数.

4.11 假设 $p: V \to \mathbb{R}$ 是非负函数使得对所有 x, α 有 $p(x\alpha) = |x|p(\alpha)$. 这确实对任何一个声称为向量长度的函数的最低要求.

(a) 相对于 p 定义连续性并证明定理 3.1 有效.

(b) 我们第二个要求是作为从 $V \times V$ 到 V 映射的加法是连续的, 而我们选定的在 0 的连续性的意思是对每个 ε 存在 δ 使得

$$p(\alpha) < \delta \quad 和 \quad p(\beta) < \delta \Rightarrow p(\alpha + \beta) < \varepsilon.$$

像在定理 3.1 中那样推理去证明存在一个常数 c 使得对所有 $\alpha, \beta \in V$ 有

$$p(\alpha + \beta) \leqslant c(p(\alpha) + p(\beta)).$$

4.12 设 V 和 W 为赋范线性空间, 并设 $f: V \times W \to \mathbb{R}$ 为有界且双线性. 设 T 为对应的从 V 到 W^* 的线性映射. 证明 T 有界并且 $\|T\|$ 是对 f 的最小界, 即最小的 b 使得对所有 α, β 有

$$|f(\alpha, \beta)| \leqslant b\|\alpha\|\ \|\beta\|.$$

4.13 设赋范线性空间 V 是一个范数直和 $M \bigoplus N$. 证明子空间 M 和 N 是 V 中闭集.(逆定理不成立)

4.14 设 N 为赋范线性空间 V 的一个闭子集. 如果 A 是倍集 $N+\alpha$, 定义 $|||A|||$ 为 $\mathrm{glb}\{\|\xi\| : \xi \in A\}$. 证明 $|||A|||$ 是商空间 V/N 上的一个范数. 再证明如果 $\overline{\xi}$ 是包含 ξ 的那个倍集, 则 $\xi \mapsto \overline{\xi}$ (V 到 V/N 上的自然投影 π) 以 1 为界.

4.15 设 V 和 W 为赋范线性空间, $T \in \mathrm{Hom}(V, W)$ 有一个包含闭子空间 N 的零空间. 证明由 $T = S \circ \pi$ 定义的从 V/N 到 W 的惟一线性映射 S(第一章定理 4.3) 是有界的且 $\|S\| = \|T\|$.

4.16 设 N 是赋范线性空间 V 的一个闭子空间, 并设 N 在纯代数意义下有一个有限维的补空间. 证明 V 是个范数直和 $M \bigoplus N$.(用上面的习题和定理 4.2 证明如果 P 是 V 沿 M 到 N 上的投影则 P 是有界的.)

4.17 设 N_1 和 N_2 是赋范线性空间 V 的闭子空间, 并设它们都具有相同的有限余维. 证明 N_1 和 N_2 为范数同构.(假定有了上面习题的结果以及第二章习题 2.11 的结果.)

4.18 证明如果 p 是向量空间 V 上的一个半范数, 则它的零集是个子空间 N, 而 p 在 N 的倍集上为常数, 并且 p 分解为 $p = q \circ \pi$, 其中 q 为 V/N 上一个范数, π 是自然投影

$\xi \mapsto \bar{\xi}: V$ 到 V/N 上. 注意, $\xi \mapsto \bar{\xi}$ 因而是从半范空间 V 到赋范空间 V/N 的一个等距满射. 等距映射是一个保持距离的映射.

3.5 无穷小

无穷小的概念在微积分的早期文献中被滥用了, 对它的描述一般说来等同于逻辑废话, 从而这个词落下了如此的坏名声以致于许多现代的书都完全回避了它. 然而它却是一个非常有用的想法, 我们把我们对微分的阐述建立在两个特殊的无穷小类的性质上; 这两个类我们称之为 "大 O" 和 "小 o"(分别以 '$\mathfrak{O}$ ' 和 '$\mathfrak{o}$ ' 表示).

最初一个无穷小被考虑为一个 "无限地小但不为 0" 的数. 当然, 没有这种数. 后来, 无穷小被考虑为一个超向于 0 为极限的一个变量. 但我们知道正是函数才有极限, 一个变量只有在某种程度上被看成是函数时才可被考虑有一个极限. 我们最终把它看成是函数 φ, 当 $t \to 0$ 时 $\varphi(t) \to 0$. 导数的定义涉及到许多这样的无穷小. 如果 $f'(x)$ 存在并取值 a, 则基本的差商 $(f(x+t)-f(x))/t$ 是两个无穷小的商, 进而 $((f(x+t)-f(x))/t)-a$ 当 $t \to 0$ 也趋于 0. 最后的这个函数在 0 没有定义, 但是我们可以回避这点, 只要我们愿意我们可以通过乘以 t 得到

$$(f(x+t)-f(x))-at=\varphi(t),$$

其中 $f(x+t)-f(x)$ 是 "f 的改变" 的无穷小, at 是线性无穷小, 而 $\varphi(t)$ 是个比 t趋向 0更快的无穷小 (即 $\varphi(t)/t$ 当 $t \to 0$ 时趋向 0). 如果我们又用 t 去除最后这个方程, 我们便看到无穷小 φ 的这个性质, 即当 $t \to 0$ 时它比 t 收敛于 0 更快, 这恰好等价于 f 的差商收敛于 a 的事实. 这便弄清楚了导数的研究包含在对无穷小变小的速率的研究之中, 而这种意释的用处很快会变得明晰起来.

定义 5.1 *赋范线性空间 V 的一个子集 A 是点 α 的一个邻域是说如果 A 包含了某个在 α 的开球. α 的一个去心邻域是 α 的一个邻域减去了点 α 自身.*

我们定义特殊的函数集 $\mathfrak{I}, \mathfrak{O}, \mathfrak{o}$ 如下, 在这些定义中我们假定了每个函数都是从赋范线性空间 V 中 0 的一个邻域到一个赋范线性空间 W 的.

$f \in \mathfrak{I}$ 是说如果 $f(0)=0$ 且 f 在 0 连续. 这些函数是无穷小.

$f \in \mathfrak{O}$ 是说如果 $f(0)=0$ 且 f 在 0 为利普希茨连续. 就是说, 存在正常数 r 和 c 使得 $\|f(\xi)\| \leqslant c\|\xi\|$ 在 $B_r(0)$ 成立.

$f \in \mathfrak{o}$ 是说如果 $f(0)=0$ 且当 $\xi \to 0$ 时 $\|f(\xi)\|/\|\xi\| \to 0$.

当空间 V 和 W 不清楚时, 我们写成 $\mathfrak{O}(V,W)$ 等来指明这点.

从 $\mathbb{R}$ 到 $\mathbb{R}$ 的一组简单的函数把这些函数类的性质差异弄得明明白白. 函数 $f(x)=|x|^{1/2}$ 在 $\mathfrak{F}(\mathbb{R},\mathbb{R})$ 中但不在 $\mathfrak{O}$ 中, $g(x)=x$ 在 $\mathfrak{O}$ 中因而也在 $\mathfrak{F}$ 中但不在 $\mathfrak{o}$ 中, 而 $h(x)=x^2$ 在所有这三个类中 (见图 3.7).

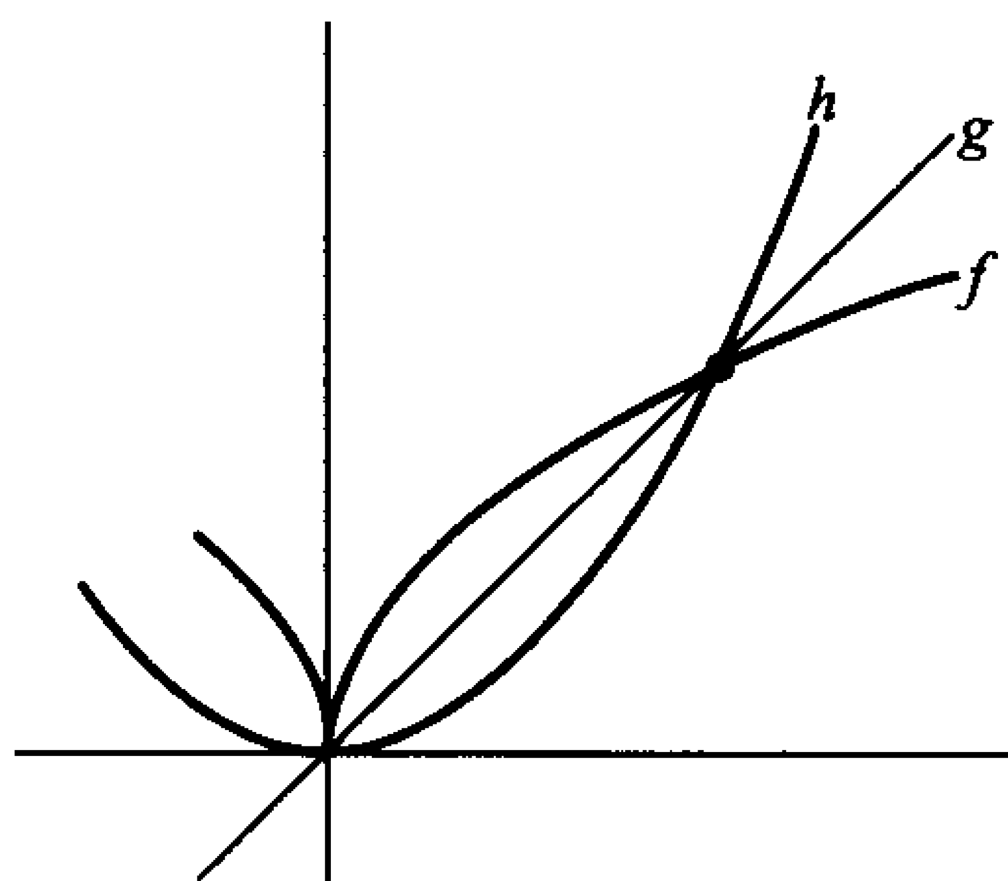


图 3.7

当 V 和 W 上的范数用等价的范数替换时 $\mathfrak{F},\mathfrak{O},\mathfrak{o}$ 仍然不变这一点是清楚的.

我们以前关于两个函数和的概念不能用两个函数 $f,g \in \mathfrak{F}(V,W)$, 这是因为它们的定义域可能各不相同. 但是 $f+g$ 定义在交集 $\mathrm{dom}f \cap \mathrm{dom}g$ 上, 这仍然是 0 的一个邻域. 那么显然 $\mathfrak{F}(V,W)$ 差不多是一个向量空间了. 惟一的麻烦出现在与等式 $f+(-f)=0$ 有关的地方; 等式左边的函数的定义域为 $\mathrm{dom}f$ 而我们自然地把 0 看作是整个 V 上的零函数.

* 消除这个难点的办法是将两个 $\mathfrak{F}$ 中函数 f 和 g 当它们在关于 0 的某个球上相同时就当作相同. 我们定义 f 与 g *等价*($f \sim g$) 当且仅当存在 0 的一个邻域, 在上面 $f=g$. 然后我们验证 (在心里) 这是一个等价关系, 从而我们现在有了一个向量空间. 它的元素被称作在 0 的*函数芽*. 严格地说, 一个芽是函数的一个等价类, 但在实用上人们总以它们的代表函数来想到芽, 这时只要记住两个函数当它们在 0 的一个邻域上一致时作为芽它们是相同的. *

正如人们可能从我们导引式的讨论中猜到的那样, 这三个类 $\mathfrak{F},\mathfrak{O},\mathfrak{o}$ 的代数性质对于微分理论来说是至关重要的. 我们把它们收集在下面的定理中.

定理 5.1 (1) $\mathfrak{o}(V,W) \subset \mathfrak{O}(V,W) \subset \mathfrak{F}(V,W)$, *并且三个类中每一个对加法和纯量乘法都是封闭的.*

(2) 如果 $f \in \mathfrak{O}(V,W)$, $g \in \mathfrak{O}(W,X)$ 则 $g \circ f \in \mathfrak{O}(V,X)$, 其中 $\mathrm{dom} g \circ f = f^{-1}[\mathrm{dom} g]$.

(3) 如果 f 或 g 如上并在 $\mathfrak{o}$ 中则 $g \circ f$ 也如此.

(4) 如果 $f \in \mathfrak{O}(V,W)$, $g \in \mathfrak{I}(V,\mathbb{R})$, 则 $fg \in \mathfrak{o}(V,W)$; 如果 $f \in \mathfrak{I}$, $g \in \mathfrak{O}$ 也是相似的.

(5) 在 (4) 中 f 和 g 中任一个在 $\mathfrak{o}$ 中而另一个仅在 0 的一个邻域中有界, 则 $fg \in \mathfrak{o}(V,W)$.

(6) $\mathrm{Hom}(V,W) \subset \mathfrak{O}(V,W)$.

(7) $\mathrm{Hom}(V,W) \cap \mathfrak{o}(V,W) = \{0\}$.

证明 设 $\mathfrak{L}_\varepsilon(V,W)$ 为所有使 $\|f(\xi)\| \leqslant \varepsilon\|\xi\|$ 在 0 的某个球成立的无穷小 f 的集合. 于是 $f \in \mathfrak{O}$ 当且仅当 f 在某个$\mathfrak{L}_\varepsilon$ 中, 而 $f \in \mathfrak{O}$ 当且仅当 f 在每个 $\mathfrak{L}_\varepsilon$ 中. 显然 $\mathfrak{o} \subset \mathfrak{O} \subset \mathfrak{I}$.

(1) 如果在 $B_t(0)$ 上 $\|f(\xi)\| \leqslant a\|\xi\|$, 在 $B_u(0)$ 上 $\|g(\xi)\| \leqslant b\|\xi\|$, 则在 $B_r(0)$ 上

$$\|f(\xi) + g(\xi)\| \leqslant (a+b)\|\xi\|,$$

其中 $r = \min\{t,u\}$. 因此 $\mathfrak{O}$ 在加法下封闭. $\mathfrak{o}$ 对加法的封闭性可相似地得到或者直接从和的极限为极限的和得出.

(2) 当 $\|\xi\| \leqslant t$ 时 $\|f(\xi)\| \leqslant a\|\xi\|$, $\|\eta\| \leqslant u$ 时 $\|g(\eta)\| \leqslant b\|\eta\|$ 则当 $\|\xi\| \leqslant t$ 和 $\|f(\xi)\| \leqslant u$, 因而当 $\|\xi\| \leqslant r = \min\{t, u/a\}$ 时有

$$\|g(f(\xi))\| \leqslant b\|f(\xi)\| \leqslant ab\|\xi\|.$$

(3) 现在假定 (2) 中的 $f \in \mathfrak{o}$. 于是给定 ε, 我们可取 $a = \varepsilon/b$ 从而当 $\|\xi\| \leqslant r$ 时

$$\|g(f(\xi))\| \leqslant \varepsilon\|\xi\|.$$

因此 $g \circ f \in \mathfrak{o}$. 对 $g \in \mathfrak{o}$ 而 $f \in \mathfrak{O}$ 的论证本质上是一样的.

(4) 在 $B_r(0)$ 上给出 $\|f(\xi)\| \leqslant c\|\xi\|$, 对给定的 ε, 我们选取 δ 使得在 $B_\delta(0)$ 上 $|g(\xi)| \leqslant \varepsilon/c$, 从而当 $\|\xi\| \leqslant \min(\delta, r)$ 时有

$$\|f(\xi)g(\xi)\| \leqslant \varepsilon\|\xi\|.$$

其他结果可相似地得到,(5) 也同样地证明.

(6) 由定义一个有界的线性变换在 $\mathfrak{O}$ 中.

(7) 假设 $f \in \mathrm{Hom}(V,W) \cap \mathfrak{o}(V,W)$. 取任一 $\alpha \neq 0$. 对给定的 ε, 选取 r 使得在 $B_r(0)$ 上 $\|f(\xi)\| \leqslant \varepsilon\|\xi\|$. 记 α 为 $\alpha = x\xi$, 其中 $\|\xi\| < r$.(求出 ξ 和 x.) 则

$$\|f(\alpha)\| = \|f(x\xi)\| = |x| \cdot \|f(\xi)\| \leqslant |x| \cdot \varepsilon \cdot \|\xi\| = \varepsilon\|\alpha\|.$$

因此对每一个正数 ε 有 $\|f(\alpha)\| \leqslant \varepsilon\|\alpha\|$, 故而 $f(\alpha) = 0$. 从而 $f = 0$,(7) 得证. □

注 f 的可加性在这里的推断中没有被用到而只用到了它的齐性. 因而由此可推导出在 $\mathfrak{o}$ 中除 $\mathfrak{O}$ 外不存在 (次数为 1 的) 齐次函数.

有时当多于一个的变量出现时有必要指明一个函数关于哪一个变量在 $\mathfrak{O}$ 或 $\mathfrak{o}$ 中. 于是我们将 "$f \in \mathfrak{O}$" 写成 "$f(\xi) = \mathfrak{O}(\xi)$", 其中 "$\mathfrak{O}(\xi)$" 用来表示 $\mathfrak{O}$ 中一个任意元素.

下面这个十分古怪的引理在我们后面对隐式定义的函数的可微性证明中要被用到. 引理隐含着 $\eta = f(\xi)$, 其中 f 是我们在研究的函数.

引理 5.1 *如果 $\eta = \mathfrak{O}(\xi) + \mathfrak{o}(\prec \xi, \eta \succ)$ 以及 $\eta = \mathfrak{F}(\xi)$, 则 $\eta = \mathfrak{O}(\xi)$.*

证明 假设条件意味着存在数 b, r_1 和 ρ 使得如果 $\|\xi\| \leqslant r_1$ 及 $\|\xi\| + \|\eta\| \leqslant \rho$ 则 $\|\eta\| \leqslant b\|\xi\| + 1/2(\|\xi\| + \|\eta\|)$; 然后, 如果 $\|\xi\|$ 小于某个 r_2 时 $\|\eta\| \leqslant \rho/2$. 因此, 如果 $\|\xi\| \leqslant r = \min\{r_1, r_2, \rho/2\}$ 则所有的条件都满足, 从而 $\|\eta\| \leqslant b\|\xi\| + 1/2(\|\xi\| + \|\eta\|)$ 成立. 但这就是不等式 $\|\eta\| \leqslant (2b+1)\|\xi\|$, 故 $\eta = \mathfrak{O}(\xi)$. □

我们必需要下面直接得到的结果.

引理 5.2 *如果 $f \in \mathfrak{O}(V,X)$, $g \in \mathfrak{O}(V,Y)$, 则 $\prec f, g \succ \in \mathfrak{O}(V, X \times Y)$. 就是说, $\prec \mathfrak{O}(\xi), \mathfrak{O}(\xi) \succ = \mathfrak{o}(\xi)$.*

证明 留给读者. □

习题

5.1 详细证明如果 V 和 W 上的范数被换成等价的范数则类 $\mathfrak{F}(V,W)$ 不变.

5.2 对 $\mathfrak{O}$ 和 $\mathfrak{o}$ 做上面同样的事.

5.3 证明 $\mathfrak{O}\mathfrak{o}$ - 定理中的 (5)(定理 5.1).

5.4 还要证明如果在 (4) 中 f 或 g 中一个在 $\mathfrak{O}$ 中而另一个只在 0 的一个邻域中有界, 则 $fg \in \mathfrak{O}(V,W)$.

5.5 证明引理 5.2.(记住, $F = \prec F_1, F_2 \succ$ 是 $F = \theta_1 \circ F_1 + \theta_2 \circ F_2$ 的不严谨用语.) 叙述对 n 个函数的推广. 叙述此定理的 θ - 形式.

5.6 给定 $F_1 \in \mathfrak{D}(V_1, W)$, $F_2 \in \mathfrak{D}(V_2, W)$, 定义从 $V = V_1 \times V_2$ 到 W 的 F 为 $F(\alpha_1, \alpha_2) = F_1(\alpha_1) + F_2(\alpha_2)$. 证明 $F \in \mathfrak{D}(V, W)$.(首先把定义方程陈述为一个涉及投影 π_1 和 π_2) 的, 但不涉及显式提及定义域向量 α_1 和 α_2 的等式.)

5.7 对给出的 $F_1 \in \mathfrak{D}(V_1, W)$ 和 $F_2 \in \mathfrak{D}(V_2, \mathbb{R})$, 准确地定义你所想到的 F_1F_2 的意思并证明它在 $\mathfrak{o}(V_1 \times V_2, W)$ 中.

5.8 定义类 $\mathfrak{D}^n$ 如下: 如果 $f \in \mathfrak{F}$ 且 $\|f(\xi)\|/\|\xi\|^n$ 在某个 0 的去心邻域中有界则 $f \in \theta^n$.(关于 α 的一个去心邻域是一个邻域减去 α.) 叙述并证明当 $f \in \mathfrak{D}^n, g \in \mathfrak{D}^m$ 时关于 $f+g$ 的一个定理.

5.9 叙述并证明当 $f \in \mathfrak{D}^n$, $g \in \mathfrak{D}^m$ 时关于 $f \circ g$ 的一个定理.

5.10 叙述并证明当 $f \in \theta^n, g \in \mathfrak{D}^m$ 时关于 fg 的一个定理.

5.11 定义一个类似的类 θ^n. 叙述并证明当 $f \in \mathfrak{D}^n, g \in \mathfrak{D}^m$ 时关于 $f \circ g$ 的一个定理.

3.6 微分

在考虑微分的概念前, 我们要复习一些由初等微积分来的几何素材. 我们这样做只是为了激发兴趣; 我们后面的理论与这个预备性的讨论是独立的.

在初等的一元微积分里函数 f 在点 a 的导数 $f'(a)$ 所具有的几何意义是 f 的图像在点 a 的切线的斜率.(当然, 根据我们的函数概念, f 的图像就是 f.) 因此这切线具有 (点斜式) 方程 $y - f(a) = f'(a)(x - a)$, 并且是仿射映射 $x \mapsto f'(a)(x - a) + f(a)$ 的图像.

通常我们考查曲线f 在点 $\prec a, f(a) \succ$ 的性质使用的是新的变量, 它们在此点都为 0. 就是说我们把一切东西都用 $s = y - f(a)$ 和 $t = x - a$ 来表达. 这个变量更换只不过是笛卡儿平面 $\mathbb{R}^2$ 中的平移 $\prec x, y \succ \mapsto \prec t, s \succ = \prec x - a, y - f(a) \succ$, 它将感兴趣的点 $\prec a, f(a) \succ$ 带到原点. 如果我们在欧几里得平面中把它画出来, 后面的图便是一个令人满意的局部模型, 于是 $\mathbb{R}^2$ 中这个平移由选择新的轴即 t 轴和 s 轴代表, 其原点在切点. 由于 $y = f(x)$ 当且仅当 $s = f(a+t) - f(a)$, 我们看出 f 在此平移下的像是函数 Δf_a, 其定义为 $\Delta f_a(t) = f(a+t) - f(a)$(见图 3.8). 当然, Δf_a 只不过是我们的老朋友, 即由 x 从 a 变到 $a + t$ 所发生在 f 中的变化.

相似地, 方程 $y - f(a) = f'(a)(x - a)$ 成为 $s = f'(a)t$, 而此切线相应地变换为一条直线即以数 $f'(a)$ 为其骨架 (矩阵) 的线性泛函 $l : t \mapsto f'(a)t$(的图像). 请

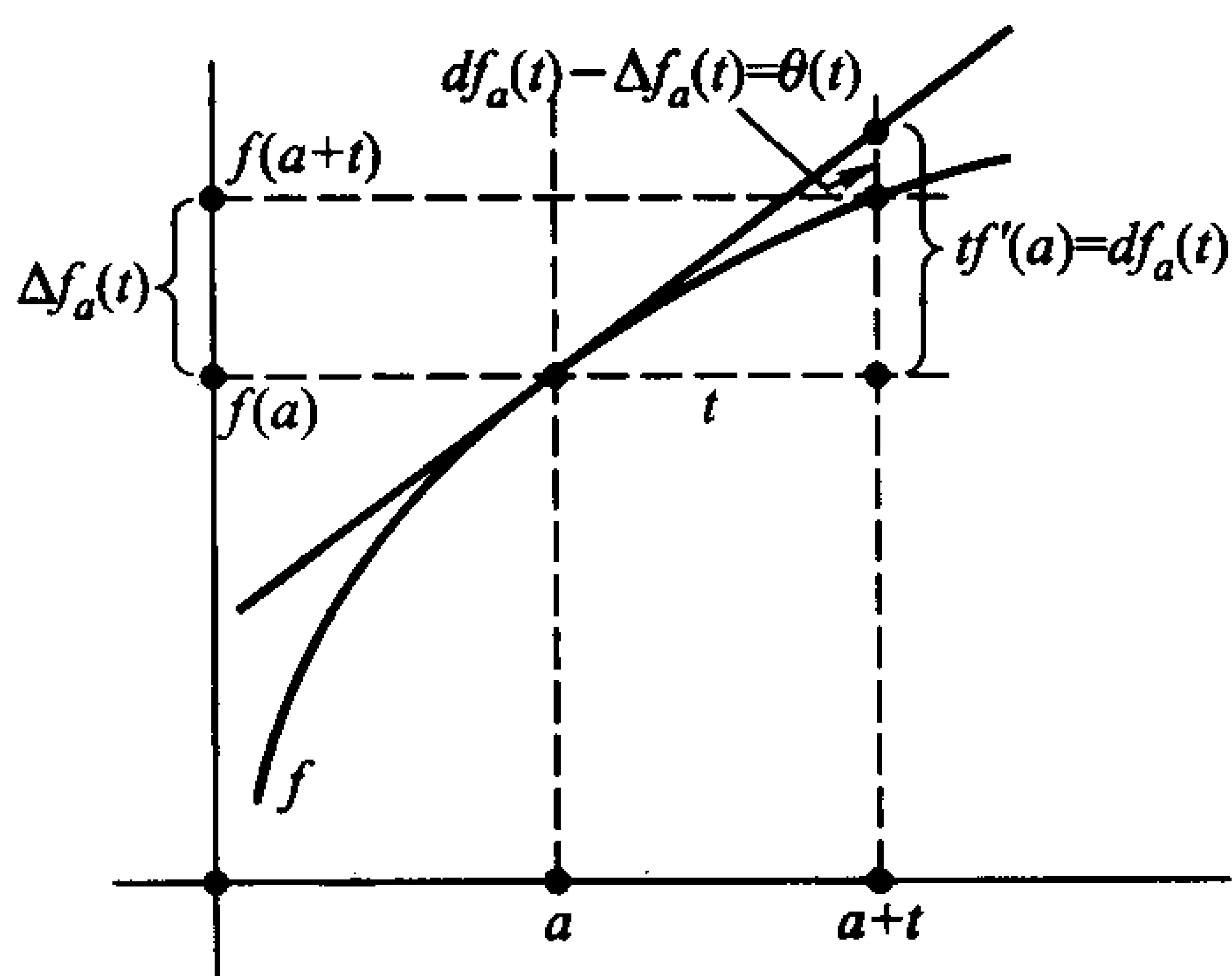


图 3.8

记住, 从在欧几里得平面中几何构图 (曲线和切线) 的观点看, 我们正在做的一切是在选取自然的坐标系, 它以切点为原点. 因此这条曲线是函数 Δf_a(的图像), 而切线是线性映射 l(的图像).

现在从 $f'(a)$ 的定义可以推导出 l 也可以被特征地描述为最紧密地近似于 Δf_a 的线性函数. 因为, 由定义,

$$\text{当} t \to 0, \frac{\Delta f_a(t)}{t} \to f'(a),$$

这与说

$$\frac{\Delta f_a(t) - l(t)}{t} \to 0 \quad \text{或} \quad \Delta f_a - l \in \mathfrak{o}$$

是完全一样的. 但是我们从 $\mathfrak{O}\mathfrak{o}$ - 定理知道把函数 Δf_a 表达为和 $l + \mathfrak{o}$ 是惟一的. 称这个惟一的线性近似为 f 在 a 的微分并以 df_a 表示. 另外, f 在 a 的微分就是那个线性函数 $l : \mathbb{R} \to \mathbb{R}$, 它在 $\Delta f_a - l\varepsilon$ 的意义下近似于 f 的实际变化 Δf_α; 我们在上面曾看到过, 如果导数 $f'(a)$ 存在, 则 f 在 a 的微分存在并以 $f'(a)$ 为其骨架 (1×1 矩阵).

相似地, 如果 f 是个二元函数, 则 f(的图像) 是笛卡儿 3 维空间 $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$ 中的曲面, 并且此曲面在 $\prec a, b, f(a,b) \succ$ 的切平面有方程 $z - f(a,b) = f_1(a,b)(x - a) + f_2(a,b)(x - b)$, 其中 $f_1 = \partial f/\partial x, f_2 = \partial f/\partial y$, 如果像上面那样, 令

$$\Delta f_{\prec a,b\succ}(s,t) = f(a+s, b+t) - f(a,b)$$

以及 $l(s,t) = sf_1(a,b) + tf_2(a,b)$, 则 $\Delta f_{\prec a,b\succ}$ 是 f 围绕着 a, b 的变化, 而 l 是具矩阵 (骨架)$\prec f_1(a,b), f_2(a,b) \succ$ 的在 $\mathbb{R}^n$ 上的线性泛函. 又, 一个标准的微积分

定理说如果 f 的偏导数连续, 则 l 以属于 $\mathfrak{o}$ 的误差近似于 $\Delta f_{\prec a,b\succ}$. 在这里 l 仍被称作 f 在 $\prec a,b\succ$ 的微分, 并以 $df_{\prec a,b\succ}$ 表示 (图 3.9). 此图中的记号已有所改变以显示 f 在 $\boldsymbol{a}=\prec a_1,a_2\succ$ 的微分 $df_{\boldsymbol{a}}$ 在 $\boldsymbol{t}=\prec t_1,t_2\succ$ 的值.

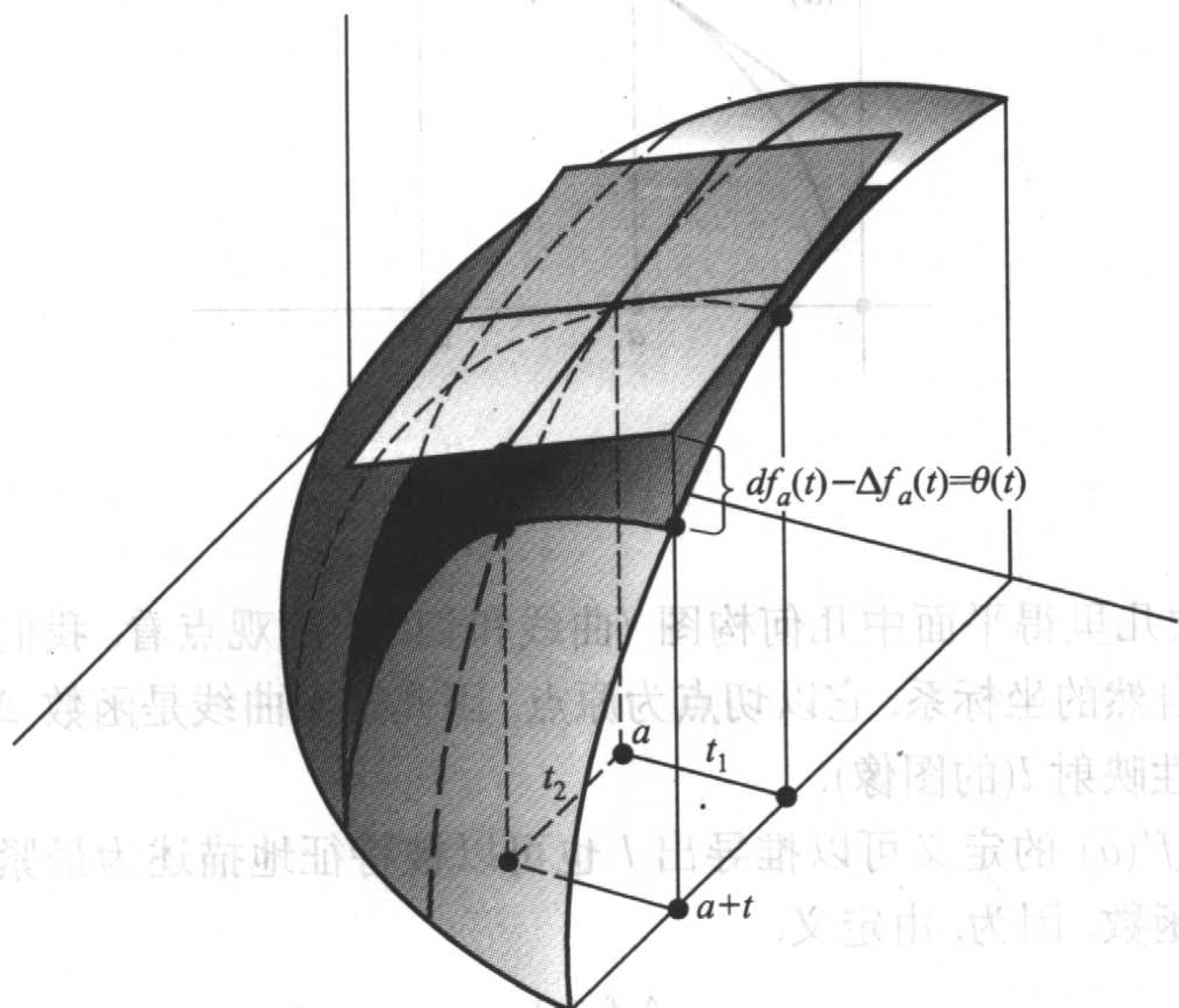


图 3.9

下面的定义现在应该是清晰的. 像上面那样, 局部函数 δF_α 被定义为 $\Delta F_\alpha(\xi)=F(\alpha+\xi)-F(\alpha)$.

定义 6.1 *设 V 和 W 为赋范线性空间, 并设 A 为 α 在 V 中的一个邻域. 映射 $F:A\to W$ 在 α 是可微的是说如果存在一个 $\mathrm{Hom}(V,W)$ 中的 T 使得 $\Delta F_\alpha(\xi)\equiv T(\xi)+\mathfrak{o}(\xi)$.*

$\mathfrak{O}\mathfrak{o}$ - 定理意味着 T 是惟一确定的, 因为如果也有 $\Delta F_\alpha=S+\mathfrak{o}$, 则 $T-S\in\mathfrak{o}$, 故由该定理的 (7) 有 $T-S=0$. 这个惟一决定的 T 被称作 *F在 α 的微分*, 并记以 dF_α. 因此 $\Delta F_\alpha=dF_\alpha+\mathfrak{o}$, 其中 dF_α 是对 ΔF_α 的那个惟一的线性近似.

* 我们的预备性讨论应该已说清楚了当 $\mathbb{R}^n$ 为定义空间时微分的定义与标准的用法一致. 但是在定义空间是无穷维空间的某些情况下, dF_α 被称为 F 在 α 的*第一变分*.这是由于虽然早期的关于变分法的作者已看出它与微分法的类比, 但他们没有认识到它是同一学科. *

我们在下两个定理中把对微分所熟知的规则收集在一起. 它们立刻由定义和 $\mathfrak{Do}$ - 定理得出.

对由 V 中 α 的邻域到 W 的所有在 α 可微的映射集合用记号 $\mathfrak{D}_\alpha(V,W)$ 表示有方便之处.

定理 6.1 (1) 如果 $F \in \mathfrak{D}_\alpha(V,W)$, 则 $\Delta F_\alpha \in \mathfrak{O}(V,W)$.

(2) 如果 $F, G \in \mathfrak{D}_\alpha(V,W)$, 则 $F+G \in \mathfrak{D}_\alpha(V,W)$ 和 $d(F+G)_\alpha = dF_\alpha + dG_\alpha$.

(3) 如果 $F \in \mathfrak{D}_\alpha(V,\mathbb{R}), G \in \mathfrak{D}_\alpha(V,W)$, 则 $FG \in \mathfrak{D}_\alpha(V,W)$ 和 $d(FG)_\alpha = F(\alpha)dG_\alpha + dF_\alpha G(\alpha)$, 第二项是个并向量.

(4) 如果 F 是 V 上的常值函数, 则 F 可微且 $dF_\alpha = 0$.

(5) 如果 $f \in \operatorname{Hom}(V,W)$, 则 F 在每个 $\alpha \in V$ 可微且 $dF_\alpha = F$.

证明 (1)$\Delta F_\alpha = dF_\alpha + \mathfrak{o} = \mathfrak{O} + \mathfrak{o} = \mathfrak{O}$, 来自 $\mathfrak{Do}$ - 定理的 (1) 和 (6).

(2) 可清楚看出 $\Delta(F+G)_\alpha = \Delta F_\alpha + \Delta G_\alpha$. 因此 $\Delta(F+G)_\alpha = (dF_\alpha + \mathfrak{o}) + (dG_\alpha + \mathfrak{o}) = (dF_\alpha + dG_\alpha) + \mathfrak{o}$, 这是由于 $\mathfrak{Do}$ - 定理的 (1). 因为 $dF_\alpha + dG_\alpha \in \operatorname{Hom}(V,W)$, 从而有 (2).

(3) $\Delta(FG)_\alpha(\xi) = F(\alpha+\xi)G(\alpha+\xi) - F(\alpha)G(\alpha)$

$$= \Delta F_\alpha(\xi)G(\alpha) + F(\alpha)\Delta G_\alpha(\xi) + \Delta F_\alpha(\xi)\Delta G_\alpha(\xi),$$

读者可用展开与消去看出以上结果. 这只是加减中间项以得到包含 Δ 的项的通常手法. 因此由 $\mathfrak{Do}$ - 定理有

$$\begin{aligned}\Delta(FG)_\alpha &= (dF_\alpha + \mathfrak{o})G(\alpha) + F(\alpha)(dG_\alpha + \mathfrak{o}) + \mathfrak{O}\mathfrak{O} \\ &= dF_\alpha G(\alpha) + F(\alpha)dG_\alpha + \mathfrak{o}.\end{aligned}$$

(4) 由 $\mathfrak{Do}$ - 定理的 (7), 如果 $\Delta F_\alpha = 0$, 则 $dF_\alpha = 0$.

(5) $\Delta F_\alpha(\xi) = F(\alpha+\xi) - F(\alpha) = F(\xi)$, 因此 $\Delta F_\alpha = F \in \operatorname{Hom}(V,W)$. □

复合函数规则多少有些复杂.

定理 6.2 如 $F \in \mathfrak{D}_\alpha(V,W), G \in \mathfrak{D}_{F(\alpha)}(W,X)$, 则 $G \circ F \in \mathfrak{D}_\alpha(V,X)$ 而且

$$d(G \circ F)_\alpha = dG_{F(\alpha)} \circ dF_\alpha.$$

证明 我们有

$$\begin{aligned}\Delta(G\circ F)_\alpha(\xi) &= G(F(\alpha+\xi)) - G(F(\alpha))\\ &= G(F(\alpha)+\Delta F_\alpha(\xi)) - G(F(\alpha))\\ &= \Delta G_{F(\alpha)}(\Delta F_\alpha(\xi))\\ &= dG_{F(\alpha)}(\Delta F_\alpha(\xi)) + \mathfrak{o}(\Delta F_\alpha(\xi))\\ &= dG_{F(\alpha)}(dF_\alpha(\xi)) + dG_{F(\alpha)}(\mathfrak{o}(\xi)) + \mathfrak{o}\circ\mathfrak{O}\\ &= (dG_{F(\alpha)}\circ dF_\alpha)(\xi) + \mathfrak{O}\circ\mathfrak{o} + \mathfrak{o}\circ\mathfrak{O}.\end{aligned}$$

因此 $\Delta(G\circ F)_\alpha = dG_{F(\alpha)}\circ dF_\alpha + \mathfrak{o}$, 并且由于 $dG_{F(\alpha)}\circ dF_\alpha \in \mathrm{Hom}(V,W)$, 定理得证. 读者应该能够对上面一串等式中每一步证明其正当性. □

习题

6.1 坐标映射 $\prec x,y\succ\mapsto x:\mathbb{R}^2\to\mathbb{R}$ 是可微的. 为什么? 它的微分是什么?

6.2 证明微分与有界线性映射的作用是可换的. 那就是说, 证明如果 $F:V\to W$ 在 α 可微而 $T\in\mathrm{Hom}(W,X)$, 则 $T\circ F$ 在 α 可微且 $d(T\circ F)_\alpha = T\circ dF_\alpha$.

6.3 证明 $F\in\mathfrak{D}_\alpha(V,\mathbb{R})$ 且 $F(\alpha)\neq 0 \Rightarrow G=1/F\in\mathfrak{D}_\alpha(V,\mathbb{R})$ 且

$$dG_\alpha = \frac{-dF_\alpha}{(F(\alpha))^2}.$$

6.4 设 $F:V\to\mathbb{R}$ 在 α 可微, 又说 $f:\mathbb{R}\to\mathbb{R}$ 是个在 $a=F(\alpha)$ 存在导数的函数. 证明 $f\circ F$ 在 α 可微且

$$d(f\circ F)_\alpha = f'(a)dF_\alpha.$$

[记住 f 在 a 的微分只是简单地乘以它的导数:$df_a(h)=hf'(a)$.] 指出前面的问题是一个特殊情形.

6.5 设 V 和 W 是赋范线性空间, 又说 $F:V\to W, G:W\to V$ 为连续映射使得 $G\circ F=I_V$, $F\circ G=I_W$. 假定 F 在 α 可微, G 在 $\beta=F(\alpha)$ 可微. 证明

$$dG_\beta = (dF_\alpha)^{-1}.$$

6.6 设 $f:V\to\mathbb{R}$ 在 α 可微. 证明 $g=f^n$ 在 α 可微且

$$dg_\alpha = nf^{n-1}(\alpha)df_\alpha.$$

(既用对乘法规则的归纳法又用复合函数规则去证明, 在第二种情形中假定 $D_x x^n = nx^{n-1}$.)

6.7 从乘积规则用归纳法证明, 如果 n 个函数 $f_i : V \to \mathbb{R}$, $i = 1, \cdots, n$ 全都在 α 可微, 则 $f = \prod_1^n f_i$ 也如此, 并且

$$df_\alpha = \sum_{i=1}^{n} \left[\prod_{j \neq i} f_j(\alpha) \right] d(f_i)_\alpha.$$

6.8 赋范线性空间 V 上的一个n 次单项式是个线性泛函的乘积 $\prod_1^n l_i (l_i \in V^*)$. 一个$n$ 次齐次多项式是 n 次单项式的有限和. 一个 n 次多项式是 i 次齐次多项式 $P_i, i = 0, \cdots, n$ 的和, 其中 P_0 为常数. 由上面的习题和其他已知的事实证明多项式处处可微.

6.9 证明如果 $F_1 : V \to W_1$ 和 $F_2 : V \to W_2$ 都在 α 可微, 则从 V 到 $W = W_1 \times W_2$ 的 $F = \prec F_1, F_2 \succ$ 也是如此 (运用单射 θ_1 和 θ_2).

6.10 不用显式计算而使用前面的习题证明由

$$\prec x, y \succ \mapsto \prec (x-y)^2, (x+y)^3 \succ$$

定义的映射 $F : \mathbb{R}^2 \to \mathbb{R}^2$ 处处可微, 现计算它在 $\prec a, b \succ$ 的微分.

6.11 设 $F : V \to X, G : W \to X$ 分别在 α 和 β 可微, 并且定义 $K : V \times W \to X$ 为

$$K(\xi, \eta) = F(\xi) + G(\eta).$$

证明 K 在 $\prec \alpha, \beta \succ$ 可微, 方法是

(a) 用直接的 Δ - 计算;

(b) 用投影 π_1 和 π_2, 以 F 和 G 来表达 K 而不直接参照变量, 然后再应用微分规则.

6.12 现假设给出了 $F : V \to \mathbb{R}$ 和 $G : W \to X$, 并定义 K 为

$$K(\xi, \eta) = F(\xi)G(\eta).$$

证明如果 F 和 G 分别在 α 和 β 可微, 则 K 在 $\prec \alpha, \beta \succ$ 可微, 证明方式按上个习题的 (b).

6.13 设 V 和 W 为赋范线性空间. 证明从 $V \times W$ 到 $\mathbb{R}$ 的映射 $\prec \alpha, \beta \succ \mapsto \|\alpha\| \|\beta\|$ 在 $\mathfrak{o}(V \times W, \mathbb{R})$ 中. 在乘积空间上使用了极大范数.

设 $f : V \times W \to \mathbb{R}$ 为有界且双线性. 这里的有界性指的是有某个 b 使 $|f(\alpha, \beta)| \leqslant b\|\alpha\| \|\beta\|$ 对所有 α, β 成立. 证明 f 处处可微并求其微分.

6.14 设 f 和 g 为 $\mathbb{R}$ 到 $\mathbb{R}$ 的可微函数. 我们从普通微积分知道复合函数规则为

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a).$$

而我们的复合函数规则是说

$$d(f \circ g)_a = df_{g(a)} \circ dg_a,$$

其中 df_x 是线性映射 $t \mapsto f'(x)t$. 证明这两个陈述是等价的.

6.15 证明 $f(x,y) = \| \prec x, y \succ \|_1 = |x| + |y|$ 在除了坐标轴上外是可微的 (即 $df_{\prec a,b \succ}$ 当 a, b 均非零时存在).

6.16 在 $\mathbb{R}^2$ 上比较 $\| \ \|_1$ 和 $\| \ \|_\infty$ 的单位球的形状, 由上面的习题猜一猜关于 $\| \ \|_\infty$ 的可微性的定理. 证明它.

6.17 设 V 和 W 是固定的赋范线性空间, 设 X_d 为 V 到 W 的所有在 0 可微的映射集合, X_0 为所有 V 到 W 的属于 $\mathfrak{o}(V,W)$ 的集合, 而 X_l 为 $\operatorname{Hom}(V,W)$. 证明 X_d 与 X_0 为向量空间且 $X_d = X_0 \bigoplus X_l$.

6.18 设 F 是个具常数 C 的利普希茨函数并在点 α 可微. 证明 $\|dF_\alpha\| \leqslant C$.

3.7 方向导数; 中值定理

方向导数构成了微分和初等微积分的导数之间的连接纽带, 并且虽然它们多加进了概念以适应于这里的格局, 但读者应该发现它们在直觉上令人满意而且在技术层面上有用.

从一个区间 $I \subset \mathbb{R}$ 到一个赋范线性空间 W 的连续函数 f 可以在完全初等微积分的意义下有一个在点 $x \in I$ 的导数 $f'(x)$:

$$f'(x) = \lim_{t \to 0} \frac{f(x+t) - f(x)}{t}.$$

这样一个函数 f 的值域是 W 中的一条曲线或弧, 并且当我们要想记住这个几何概念时, 按习惯称 f 自身为一条参数 (化的) 弧. 如果 $f'(x)$ 存在, 我们也将称它为弧 f 在 x 的切向量.这个术语正像图 3.10 所表示的, 符合于我们的几何直观. 为简明起见我们已令 $x = 0$ 及 $f(0) = 0$. 如果 $f'(x)$ 存在, 我们则说参数弧f 在

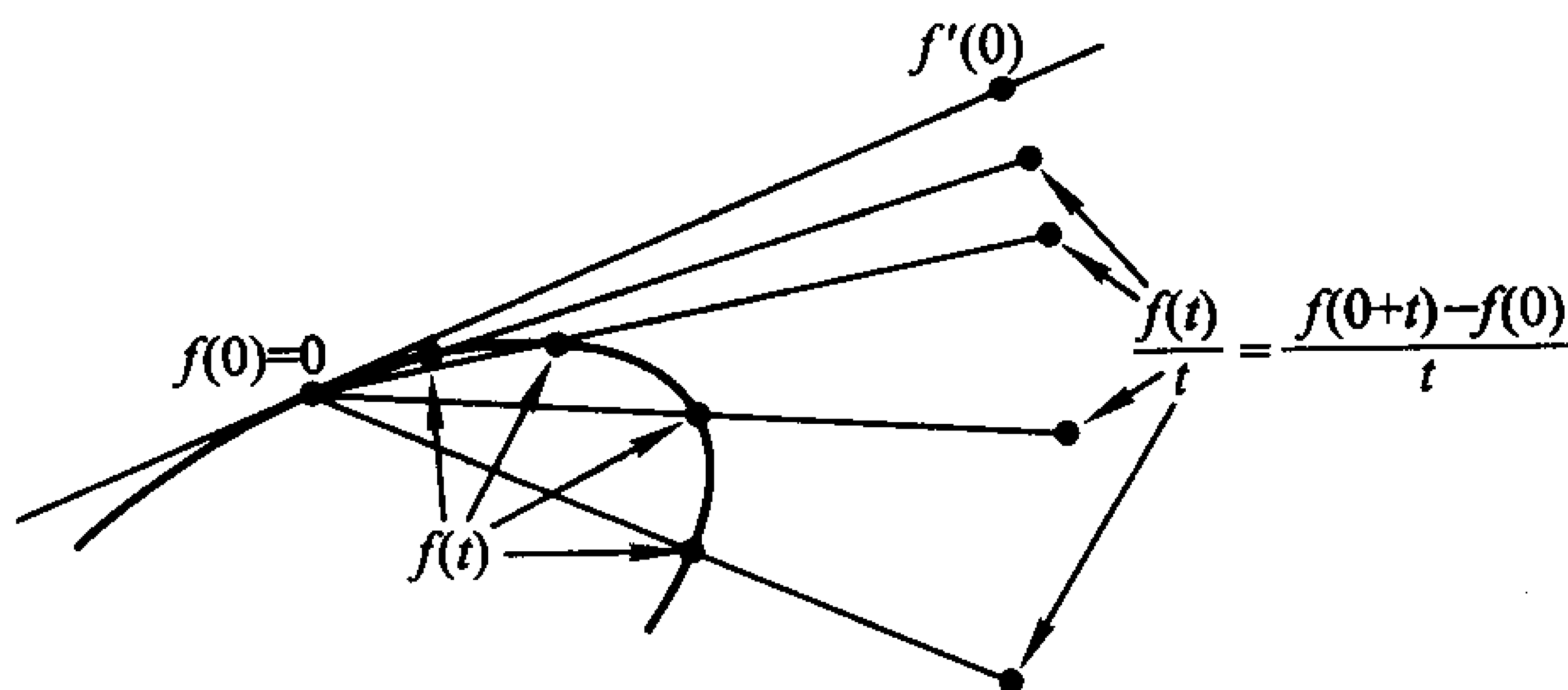


图 3.10

x光滑. 我们也说 f 在 $\alpha = f(x)$ 光滑, 但是如果 f 不是单射的 (即如果此弧自交) 则这个术语便是含糊不清的. 一条弧为光滑是说它在每个参数值上光滑.

我们自然想要知道切向量 $f'(x)$ 的存在性与 f 在 x 的可微性之间的关系. 如果 df_x 存在, 则作为 $\mathbb{R}$ 上的一个线性映射, 它仅仅是 "用" 固定的向量 α 即它的骨架去乘, $df_x(h) = hdf_x(1) = h\alpha$, 从而我们预料 α 就是切向量 $f'(x)$. 我们来证明它和对在第 6 节中我们预备性讨论中的普通微积来说的逆结果. 实际上, 我们的论证对向量 (值) 函数是有效的, 但不管怎样我们还将重复它.

当我们把一个实变量的向量 (值) 函数想成是一段弧时, 我们常常用希腊字母像 'λ' 和 'γ' 代表这个函数, 我们在下面就这样做. 当然这无论如何没有改变被证明的东西, 只不过稍稍使人联想到一种几何的解释.

定理 7.1 *一个参数弧 $\gamma : [a, b] \to V$ 在 $x \in (a, b)$ 可微当且仅当切向量 (导数)$\alpha = \gamma'(x)$ 存在, 这时此切向量是微分 $d\gamma_x(h) = h\gamma'(x) = h\alpha$ 的骨架.*

证明 如果参数弧 $\gamma : [a, b] \to V$ 在 $x \in (a, b)$ 可微, 则 $d\gamma_x(h) = hd\gamma_x(1) = h\alpha$, 其中 $\alpha = d\gamma_x(1)$. 由于 $\Delta\gamma_x - d\gamma_x \in \mathfrak{o}$, 从而给出了 $\|\Delta\gamma_x(h) - h\alpha\|/|h|$ 当 $h \to 0$ 时趋向于 0, 故 $\Delta\gamma_x(h)/h \to \alpha$. 因此 α 是在通常意义下的导数 $\gamma'(x)$. 把上面的步骤反过来我们看到 $\gamma'(x)$ 的存在性意味着 γ 在 x 的可微性. □

现在设 F 为赋范线性空间 V 中一个开集 A 到赋范线性空间 W 的一个函数. 研究在 A 中一点 α 的邻域中 F 的行为的一种办法是考虑它在每条经过 α 的直线上的行为. 那就是说, 我们暂时把 F 限制于一维区域上来研究它. 这样做的好处是被限制了的 F 便简单地成为一条参数弧, 而它的微分则简单地是乘以它的通常的导数.

对任意非零的 $\xi \in V$, 通过 α 沿方向 ξ 的直线具有参数表示 $t \mapsto \alpha + t\xi$. F 在此直线上的限制是参数弧 $\gamma : \gamma(t) = F(\alpha + t\xi)$. 它在原点 $t = 0$ 的切向量 (导数), 如果存在, 被称为*F 在 α 沿方向 ξ 的导数*, 或者*F 在 α 的关于 ξ 的导数*, 记其为 $D_\xi F(\alpha)$. 显见,

$$D_\xi F(\alpha) = \lim_{t \to 0} \frac{F(\alpha + t\xi) - F(\alpha)}{t}.$$

将此与我们原来 f' 的定义比较, 我们看出参数弧 γ 的切向量 $\gamma'(x)$ 是关于 $\mathbb{R}$ 中标准的基向量 1 的方向导数 $D_1\gamma(x)$.

严格地说, 我们错误地使用了 "方向" 这个词, 因为不同的向量可以有相同的方向. 因而如果 $\eta = c\xi, c > 0$, 则 η 和 ξ 指向同一方向, 但是因为 $D_\xi F(\alpha)$ 对

ξ 是线性的 (如同我们马上会看到的那样), 它们相伴的导数是不同的:$D_\eta F(\alpha) = cD_\xi F(\alpha)$.

现在我们要建立方向导数与微分之间的关系, 前者是向量, 所者是线性映射. 我们在上面看到过, 对一条弧 γ 而言可微性等价于 $\gamma'(x) = D_1\gamma(x)$ 的存在性. 在一般情况下这种关系并不像对一条弧时那样简单, 但沿一个方向一切都可顺利进行.

定理 7.2 *如果 F 在 α 可微, 并且如果 λ 是任一条通过 α 的光滑弧线, 满足 $\alpha = \lambda(x)$, 则 $\gamma = F \circ \lambda$ 在 x 光滑, 且 $\gamma'(x) = dF_\alpha(\lambda'(x))$. 特别地, 如果 F 在 α 可微, 则所有的方向导数 $D_\xi F(\alpha)$ 存在并且 $D_\xi F(\alpha) = dF_\alpha(\xi)$.*

证明 γ 的光滑性等价于它在 x 的可微性, 因而由复合函数定理可以得到. 又, $\gamma'(x) = d\gamma_x(1) = d(F \circ \lambda)_x(1) = dF_\alpha(d\lambda_x(1)) = dF_\alpha(\lambda'(x))$. 如果 λ 是参数直线 $\lambda(t) = \alpha + t\xi$, 则它具有常数导数 ξ, 而且由于在这里 $\alpha = \lambda(0)$, 那么上面的公式成为 $\gamma'(0) = dF_\alpha(\xi)$. 就是说, $D_\xi F(\alpha) = \gamma(0)' = dF_\alpha(\xi)$. □

反过来, 一个函数 F 在一点 α 的所有的方向导数 $D_\xi F(\alpha)$ 的存在意味着 F 在 α 的可微性是不正确的. 最容易的反例涉及到齐次函数的概念. 一个函数 $F: V \to W$ 是*齐次的*是说如果对所有 x 和 ξ 有 $F(x\xi) = xF(\xi)$. 对这样一个函数因为弧线 $\gamma(t) = F(0 + t\xi) = tF(\xi)$ 是线性的, 且 $\gamma'(0) = F(\xi)$. 因此*一个齐次函数 F 在 0 的所有方向导数均存在*, 且 $D_\xi F(0) = F(\xi)$. 如果 F 也在 0 可微, 则 $dF_0(\xi) = D_\xi F(0) = F(\xi)$, 从而 $F = dF_0$. *因此一个可微的齐次函数必是线性的.* 所以任一非线性齐次函数 F 是一个使得对所有 $\xi, D_\xi F(0)$ 存在而 dF_0 不存在的函数. 取最简单的可能的情况: 定义 $F: \mathbb{R}^2 \to \mathbb{R}$ 为如果 $\prec x, y \succ \neq \prec 0, 0 \succ$ 时 $F(x, y) = x^3/(x^2 + y^2)$, 而 $F(0, 0) = 0$, 则

$$F(tx, ty) = tF(x, y),$$

故而 F 是齐次的但不是线性的.

但是, 如果 V 为有限维, 并且如果对张成空间的向量集合中的每个 ξ, 方向导数 $D_\xi F(\alpha)$ 存在, 同时是在一个开集 A 上的 α 的连续函数, 则在 A 上 F 是连续可微的. 对此事实的证明依赖于我们接下来要处理的中值定理, 而要完成此证明则要到第九章 (定理 9.3).

读者是把中值定理作为微积分的一个奠基性定理而记住的, 在我们的一般理论中恰好也是如此. 我们要在下一节中应用它去证明上面提及的定理的一个

一般形式, 实际上我们更为高级的所有工作都有赖于它. 在这里通常的中值定理没有完全恰当的类比. 取而代之, 我们将证明一个定理, 它在单变量微积分中是中值定理的一个简单推论.

定理 7.3 设 f 为从闭区间 $[a,b]$ 到一个赋范空间的一个连续函数 (参数弧), 并假设 $f'(t)$ 存在且对所有 $t \in (a,b)$ 满足 $\|f'(t)\| \leqslant m$. 则 $\|f(b) - f(a)\| \leqslant m(b-a)$.

证明 固定一个 $\varepsilon > 0$, 并设 A 为 $[a,b]$ 中满足

$$\|f(x) - f(a)\| \leqslant (m+\varepsilon)(x-a) + \varepsilon.$$

的点 x 的集合. 因为 f 在 a 为连续故 A 至少包含了一个小区间 $[a,c]$. 设 $l = \text{lub}A$. 于是由 f 在 l 的连续性有 $\|f(l)-f(a)\| \leqslant (m+\varepsilon)(l-a)+\varepsilon$. 因此 $l \in A$ 且 $a < l \leqslant b$. 我们断言 $l = b$. 因为如果 $l < b$, 则 $f'(l)$ 存在且 $\|f'(l)\| \leqslant m$. 因此存在一个 δ 使得当 $|x-l| \leqslant \delta$ 时

$$\|[f(x) - f(l)]/(x-l)\| < m + \varepsilon.$$

由此得到

$$\begin{aligned}\|f(l+\delta) - f(a)\| &\leqslant \|f(l+\delta) - f(l)\| + \|f(l) - f(a)\| \\ &\leqslant (m+\varepsilon)\delta + (m+\varepsilon)(l-a) + \varepsilon \\ &\leqslant (m+\varepsilon)(l+\delta-a) + \varepsilon,\end{aligned}$$

故 $l + \delta \in A$, 矛盾. 因此 $l = b$. 我们于是有

$$\|f(b) - f(a)\| \leqslant (m+\varepsilon)(b-a) + \varepsilon,$$

并且因为 ε 是任意的, 故 $\|f(b) - f(a)\| \leqslant m(b-a)$. □

下面更为一般的中值定理具有通常使用的形式. 按惯例, F 和 G 是从 V 的一个子集到 W 的映射.

定理 7.4 如果 F 在球 $B_r(\alpha)$ 上可微, 且如果对此球内每个 β 有 $\|dF_\beta\| \leqslant \varepsilon$, 则只要 β 和 $\beta+\xi$ 在此球内就有 $\|\Delta F_\beta(\xi)\| \leqslant \varepsilon\|\xi\|$. 更一般地, 如果球 $B_r(\alpha)$ 被换为任一凸集C, 则相同的结果成立.

证明 从 β 到 $\beta+\xi$ 的线段是参数弧 $\lambda(t) = \beta+t\xi$ 在 $[0,1]$ 到 V 的值域. 如果 β 和 $\beta+\xi$ 在球 $\beta_r(\alpha)$ 中, 则此线段是此球的一个子集. 令 $\gamma(t) = F(\beta+t\xi)$, 我们便

有 $\gamma'(x) = dF_{\beta+x\xi}(\lambda'(x)) = dF_{\beta+x\xi}(\xi)$, 这可由定理 7.2 得到. 因此 $\|\gamma'(x)\| \leqslant \varepsilon\|\xi\|$ 在 [0,1] 上成立. 于是中值定理意味着

$$\begin{aligned}\|\Delta F_\beta(\xi)\| = \|F(\beta+\varepsilon) - F(\beta)\| &= \|\gamma(1) - \gamma(0)\| \\ &\leqslant \varepsilon\|\xi\|(1-0) = \varepsilon\|\xi\|,\end{aligned}$$

这是所要的不等式. 我们所用到的 $B_r(\alpha)$ 的惟一性质是它包含了联结它中任意两点的线段. 这就是凸性的定义, 故此定理对任意凸集为真. □

推论 7.1 如果 G 在凸集 C 上可微, 如果 $T \in \mathrm{Hom}(V,W)$, 又如果对所有 C 中的 β 有 $\|dG_\beta - T\| \leqslant \varepsilon$, 则只要 β 和 $\beta+\xi$ 在 C 中便有 $\|\Delta G_\beta(\xi) - T(\xi)\| \leqslant \varepsilon\|\xi\|$.

证明 令 $F = G - T$ 且注意到有 $dF_\beta = dG_\beta - T$ 和 $\Delta F_\beta = \Delta G_\beta - T$ 即可. □

我们说几句关于记号的话来结束本节. 留意一下在等式 $(D_\xi F)(\alpha) = dF_\alpha(\xi)$ 中变量位置的颠倒. 这个差别具有实际的重要性. 我们有一个二个变量 'α' 和 'ξ' 的函数, 我们可以通过固定一个变量把它转换为另一个变量的单变量函数; 从技术层面上说, 把那个固定的变量放在下标位置有其方便之处. 因此我们考虑 $dF_\alpha(\xi)$ 为 α 固定并有 $\mathrm{Hom}(V,W)$ 中函数 dF_α, 而在 $(D_\xi F)(\alpha)$ 中我们则取 ξ 固定并有方向导数 $D_\xi F : A \to W$, 它是沿固定方向的一个 α 的函数, 是对作为 $\boldsymbol{a}$ 的函数的任意普通偏导数 $\partial F/\partial x_i(\boldsymbol{a})$ 的推广记号. 我们也能够把这种变量的下标位置的含意用点记号 (0.10 节) 来表示: 当我们写下 $D_\xi F(\alpha)$ 时我们想到的是函数 $D_\xi F(\cdot)$ 在 α 的值.

在后面的章节中我们还要用到第三种记号, 它把函数符号放在下标位置. 我们记

$$J_F(\alpha) = dF_\alpha.$$

这个记号表示在某个讨论中映射 F 要被固定, 并放在下标位置以把它 "晾在一边".

如果 F 在开集 A 中每点可微, 我们自然会考虑作为 A 到 $\mathrm{Hom}(V,W)$ 的函数 $dF : \alpha \mapsto dF_\alpha$. 用 "$J$" 记号即 $dF = J_F$. 本章后面我们要考虑这个映射在 α 的可微性. 这个二阶微分$d^2F_\alpha = d(dF)_\alpha$ 的概念最初看来或许是令人困惑的, 因而现在预先看一看会让以后的讨论容易一些. 我们不过是有了一个从赋范线性空间 V 中开集 A 到赋范线性空间 $X = \mathrm{Hom}(V,W)$ 的一个新的映射 $G = dF$, 并且我们在考虑它在 α 的可微性. 回过来参看第一章的定理 6.1, 我们知道了

$dG_\alpha = d^2F_\alpha$ 由对偶性等价于一个双线性映射 ω, 它从 $V\times V$ 到 W: 由于 $dG_\alpha(\xi)$ 自身是 $\mathrm{Hom}(V,W)$ 中的一个变换, 我们可将其在 η 上赋值, 并定义 ω 为

$$\omega(\xi,\eta) = (dG_\alpha(\xi))(\eta).$$

用点记号在这里也是会有帮助的. 映射 $\alpha\mapsto dF_\alpha$ 简单地记为 $dF_{(\cdot)}$, 而 G 由 $G(\cdot)=dF_{(\cdot)}$ 定义. 而后, $dG_\alpha(\xi)$ 是个映射的事实可写成 $dG_\alpha(\xi)(\cdot)$ 来予以强调. 在这里的每一种情形中我们均有一个单变量的函数, 点只是提醒我们这个事实并告诉我们当要赋值时该把变量放在哪里. 在 ω 的情形则是点记号的原来的用法. 即 $\omega(\xi,\cdot)=dG_\alpha(\xi)$.

习题

7.1 给出 $f:\mathbb{R}\to\mathbb{R}$ 使 $f'(a)$ 存在, 用差商极限的直接取值证明 "方向导数"$D_bf(a)$ 为 $bf'(a)$.

7.2 设 f 是 n 维空间 V 上的实函数, 并假定 f 在 $\alpha\in V$ 可微: 证明使导数 $D_\xi F(\alpha)$ 为 0 的方向 ξ 构成 V 的一个 $(n-1)$ 维子空间 (或是整个 V). 如果 f 映射 V 到一个二维空间 W 能得什么相似的结论?

7.3 (a) 用关于极限的直接论证证明, 如果 f 和 g 是从区间 $I\subset\mathbb{R}$ 到赋范线性空间 V 的两个函数, 并且如果 $f'(x)$ 和 $g'(x)$ 都存在, 则 $(f+g)'(x)$ 存在且 $(f+g)'(x)=f'(x)+g'(x)$.
(b) 作为定理 7.1,7.2 和第 6 节的微分规则的推论证明同一结果.

7.4 (a) 给出 $f:I\to V$ 和 $g:I\to W$, 用直接的极限论证证明如果 $f'(x)$, $g'(x)$ 都存在, 并且如果 $F=\prec f,g\succ: I\to V\times W$, 则 $F'(x)$ 存在且 $F'(x)=\prec f'(x),g'(x)\succ$.
(b) 由定理 7.1,7.2 和第 6 节的微分规则并利用关系 $F=\theta_1\circ f+\theta_2\circ g$ 证明同一结果.

7.5 按上面两个习题的精神, 对弧的导数叙述一个乘积规则并按上面 (b) 的证明方式证明它.

7.6 求在 $t=0$ 和 $t=\pi/2$ 时弧 $\prec e^t,\sin t\succ$ 的切向量.[用习题 7.4(a)] 什么是在这两个点上此系数弧的微分? 即如果 $f(t)=\prec e^t,\sin t\succ$, 什么是 df_0 和 $df_{\pi/2}$?

7.7 设 $F:\mathbb{R}^2\to\mathbb{R}^2$ 为映射 $\prec x,y\succ\mapsto\prec 3x^2y, x^2y^3\succ$. 计算方向导数 $D_{\prec 1,2\succ}F(3,-1)$:
(a) 作为弧 $f\circ\lambda$ 在 $\prec 3,-1\succ$ 的切向量, 其中 λ 是经过 $\prec 3,-1\succ$, 方向为 $\prec 1,2\succ$ 的直线;
(b) 先计算 $dF_{\prec 3,1\succ}$, 然后再在 $\prec 1,2\succ$ 计值.

7.8 设 λ 和 μ 为向量空间 V 上任意两个泛函. 沿直线 $\xi=t\alpha$ 有乘积 $f(\xi)=\lambda(\xi)\mu(\xi)$; 算出 $D_\alpha f(\alpha)$. 现沿一般直线 $\xi=t\alpha+\beta$ 对 f 赋值并由此计算 $D_\alpha f(\beta)$.

7.9 由计算微分的方法解上面的习题.

7.10 如果 $f:\mathbb{R}^n \to \mathbb{R}$ 在 $\boldsymbol{a}$ 可微, 我们知道作为 $\mathbb{R}^n$ 上线性泛函的 $df_{\boldsymbol{a}}$ 由它的骨架 n 联组 $\boldsymbol{L}$ 按下面公式给出:

$$df_{\boldsymbol{a}}(x) = (\boldsymbol{L}, \boldsymbol{x}) = \sum_1^n l_i x_i.$$

在此情形中我们称这个 n 联组 $\boldsymbol{L}$ 为 f 在 $\boldsymbol{a}$ 的梯度. 由施瓦兹不等式 (习题 2.3) 证明如果我们用欧几里得长为 1 的向量 $\boldsymbol{y}$, 则方向导数 $D_{\boldsymbol{y}}f(\boldsymbol{a})$ 当 $\boldsymbol{y}$ 指向 f 的梯度方向时取得极大值.

7.11 设 W 为赋范线性空间, V 为满足 $\lambda(0) = 0, \lambda'(0)$ 存在的参数弧 $\lambda : [-1, 1] \to W$ 的集合. 证明 V 是个向量空间, 而且 $\lambda \mapsto \lambda'(0)$ 是由 V 到 W 满射的线性映射. 用文字描述商空间 V/N 中的元素, 其中 N 为上面映射的零空间.

7.12 找出另一个齐次非线性函数. 计算它的方向导数 $D_\xi F(0)$, 它再次表明它们没有造出一个线性映射.

7.13 证明如果 F 是从赋范线性空间 V 中开球 B 到赋范线性空间 W 的可微映射, 且它对 B 中所有的 α 有 $dF_\alpha = 0$, 则 F 是一个常值函数.

7.14 把上面的习题推广到这种情形:F 的定义域是个开集 A, 而 A 中任意两个点可用 A 中的一条光滑弧线联结的性质. 用反例证明此结果不能推广到 F 的定义域 A 为任意开集的情形.

7.15 证明下面中值定理的推广. 设 f 是从闭区间 $[a, b]$ 到赋范线性空间 V 的连续映射, 又设 g 为 $[a, b]$ 上一个连续的实函数. 假如对开区间 (a, b) 中每点 $f'(t)$ 和 $g'(t)$ 都存在, 并且在 (a, b) 上 $\|f'(t)\| \leqslant g'(t)$. 于是

$$\|f(b) - f(a)\| \leqslant g(b) - g(a).$$

[考虑使得 $\|f(x) - f(a)\| \leqslant g(x) - g(a) + \varepsilon(x - a) + \varepsilon$ 的点 x.]

3.8 微分与积空间

本节我们要把微分规则与作为有限笛卡儿乘积的向量空间所产生的特殊结构联系起来. 当我们处理值域时, 这种考虑是微不足道的, 然而当定义域是乘积空间时, 我们便与一个深刻的定理直接有关了. 这些一般性的乘积考量在下一节将专门限制在 $\mathbb{R}^n$ 空间上, 但它们确有更广泛的用途, 这会在本章后面几节及本书后面几章中看到.

我们知道一个 m 联组的具有公共定义域的函数 $F^i : A \to W_i, i = 1, \cdots, m$ 等价于一个单一的 m 联组取值的函数

$$F : A \to W = \prod_1^m W_i,$$

$F(\alpha)$ 为 m 元组 $\{F^i(\alpha)\}_1^m$, 其中每个 $\alpha \in A$. 我们现在来验证那个显然而必然的事实, 即 F 在 α 可微当且仅当每个 F^i 在 α 可微.

定理 8.1 *给出 $F^i : A \to W_i, i = 1, \cdots, m, F = \prec F^1, \cdots, F^m \succ$, 则 F 在 α 可微当且仅当所有的函数 F^i 如此, 此时有*

$$dF_\alpha = \prec dF_\alpha^1, \cdots, dF_\alpha^m \succ .$$

证明 严格地说, $F = \sum_1^m \theta_i \circ F^i$, 其中 θ_j 为 W_i 到乘积空间 $W = \prod_1^m W_i$ 的内射 (见 1.3 节). 由于每个 θ_i 是线性的, 故而是可微的, 且 $d(\theta_i)_\alpha = \theta_i$, 因此我们看到如果每个 F^i 在 α 可微则 F 也如此, 并且 $dF_\alpha = \sum_1^m \theta_i \circ dF_\alpha^i$. 不严格地说, 此即 $dF_\alpha = \prec dF_\alpha^1, \cdots, dF_\alpha^m \succ$. 反方向的结论相似地由 $F^i = \pi_i \circ F$ 得到, 其中 π_j 是 $\prod_1^m W_i$ 到 W_j 的投影. □

定理 8.1 和 7.1 有下面显见的推论 (它也可用涉及到的极限容易地得到证明).

引理 8.1 *如果 f_i 为 $[a, b]$ 到 W_i 的弧, $i = 1, \cdots, n$, 而 f 是 n 联组值的弧 $f = \prec f_1, \cdots, f_n \succ$, 则 $f'(x)$ 存在当且仅当 $f_i'(x)$ 对每个 i 都存在, 这时 $f'(x) = \prec f_1'(x), \cdots, f_n'(x) \succ$.*

当定义域空间 V 是一个积空间 $\prod_1^n V_j$ 时情形则更加复杂. n 个向量变量的函数 $F(\xi_1, \cdots, \xi_n)$ 没有分解为一个等价的 n 联组的函数. 进而, 虽然它的微分 dF_α 分解为偏微分 $\{dF_\alpha^j\}$ 的等价 n 联组, 我们却没有简单的定理, 即 dF_α 存在当且仅当所有偏导数 dF_α^j 存在.

当然, 我们把一个 n 个向量变量的函数看作是单个 n - 联组变量 $\boldsymbol{\xi} = \prec \xi_1, \cdots, \xi_n \succ$ 的函数, 故而当我们考虑 F 的可微性时原则上没有什么新东西. 然而, 当我们考虑复合 $F \circ G$ 时, 其内层函数 G 现在必定是一个取值 n - 联组的函数 $G = \prec g^1, \cdots, g^n \succ$, 其中 g^i 是从某个赋范线性空间 X 的一个开子集 A 到 V_i, 从而我们自然试图把 $F \circ G$ 的微分用微分 dg^i 来表达. 要做到这个我们需要 F 的*偏导数* dF_α^i. 我们暂时定义 F 在 $\boldsymbol{\alpha} = \prec \alpha_1, \cdots, \alpha_n \succ$ 的第 j 个偏导数

为微分 $dF_{\boldsymbol{\alpha}}$ 在 V_i 上的限制, 这时我们将 V_j 考虑为 $V=\prod_1^n V_i$ 的子空间. 如通常那样, 这确实与 V_j 到 $\prod_1^n V_i$ 的内射有直接关系, 因而我们形式的 (临时) 定义为

$$dF_{\boldsymbol{\alpha}}^j = dF_{\boldsymbol{\alpha}} \circ \theta_j.$$

那么, 由于 $\boldsymbol{\xi} = \prec \xi_1, \cdots, \xi_n \succ = \sum_1^n \theta_i(\xi_i)$, 我们有

$$dF_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_1^n dF_{\boldsymbol{\alpha}}^i(\xi_i).$$

相似地, 由于 $G = \prec g^1, \cdots, g^n \succ = \sum_1^n \theta_i \circ g^i$, 我们有

$$d(F \circ G)_\gamma = \sum_1^n dF_{G(\gamma)}^i \circ dg_\gamma^i,$$

我们称其为*一般链规则*.在此公式中的 "i" 上标有含混的地方: 为了更准确我们应写成 $(dF)_{\boldsymbol{\alpha}}^i$ 和 $d(g^i)_\gamma$.

我们将逐步转到偏微分的真正定义. 由于

$$\Delta F_{\boldsymbol{\alpha}} \circ \theta_j = (dF_{\boldsymbol{\alpha}} + \mathfrak{o}) \circ \theta_j = dF_{\boldsymbol{\alpha}} \circ \theta_j + \mathfrak{o} = dF_{\boldsymbol{\alpha}}^j + \mathfrak{o},$$

我们看出 $dF_{\boldsymbol{\alpha}}^i$ 能够被直接地描绘如下面与 $dF_{\boldsymbol{\alpha}}$ 无关:

$dF_{\boldsymbol{\alpha}}^i$ 是 $\mathrm{Hom}(V_i, W)$ 中惟一的元素 T_i 使得 $\Delta F_{\boldsymbol{\alpha}} \circ \theta_i = T_i + \mathfrak{o}$.

那就是说, $dF_{\boldsymbol{\alpha}}^i$ 是将 $F(\xi_1, \cdots, \xi_n)$ 中除 ξ_i 的其他变量固定于值 $\xi_j = \alpha_j$ 得到的单变量 ξ_i 的函数在 α_i 的微分. 因为在实用中作为主要现象我们遇到的常常是这种可偏微分性, 故这是重要的. 因此我们将这个直接描述作为 dF_α^i 的定义, 在上面随其后面的启发性计算是下面引理的证明.

引理 8.2 *如果 A 是积空间 $V=\prod_1^n V_i$ 的开子集, 又如果 $F: A \to W$ 在 $\boldsymbol{\alpha}$ 可微, 则所有的偏微分 $dF_{\boldsymbol{\alpha}}^i$ 存在, 且 $dF_{\boldsymbol{\alpha}}^i = dF_{\boldsymbol{\alpha}} \circ \theta_i$.*

于是是否所有偏微分 $dF_{\boldsymbol{\alpha}}^i$ 的存在性意味着 $dF_{\boldsymbol{\alpha}}$ 的存在性的问题出现了. 一般来说答案是否定的, 我们将在下一节中看到这一点; 然而如果对一个开集 A 中的每个 $\boldsymbol{\alpha}$ 所有的偏微分 $dF_{\boldsymbol{\alpha}}^i$ 存在并且是 $\boldsymbol{\alpha}$ 的连续函数, 则 F 在 A 上连续可微. 注意引理 8.2 和投影 - 内射恒等式向我们表明如果 $dF_{\boldsymbol{\alpha}}$ 存在它会是什么样子: $dF_{\boldsymbol{\alpha}}^i = dF_{\boldsymbol{\alpha}} \circ \theta_i$ 和 $\sum \theta_i \circ \pi_i = I$ 一起给出 $dF_{\boldsymbol{\alpha}} = \sum dF_{\boldsymbol{\alpha}}^i \circ \pi_i$.

定理 8.2 设 A 为赋范线性空间 $V = V_1 \times V_2$ 的一个开子集, 并假定 $F: A \to W$ 在 A 上有连续的偏微分 $dF^1_{\prec\alpha,\beta\succ}$, $dF^2_{\prec\alpha,\beta\succ}$. 于是 $dF_{\prec\alpha,\beta\succ}$ 在 A 上存在且连续, 且

$$dF_{\prec\alpha,\beta\succ}(\xi,\eta) = dF^1_{\prec\alpha,\beta\succ}(\xi) + dF^2_{\prec\alpha,\beta\succ}(\eta).$$

证明 我们在 $V = V_1 \times V_2$ 上使用和范数. 给出 ε, 我们选取 δ 使对在 $\prec \alpha, \beta \succ$ 的 δ - 球中的每个 $\prec \mu, \nu \succ$ 有 $\|dF^i_{\prec\mu,\nu\succ} - dF^i_{\prec\alpha,\beta\succ}\| < \varepsilon$, $i = 1, 2$. 令

$$G(\xi) = F(\alpha+\xi, \beta+\eta) - dF^1_{\prec\alpha,\beta\succ}(\xi),$$

我们有当 $\| \prec \xi, \eta \succ \| < \delta$ 时

$$\|dG_\xi\| = \|dF^1_{\prec\alpha+\xi,\beta+\eta\succ} - dF^1_{\prec\alpha,\beta\succ}\| < \varepsilon,$$

而定理 7.4 的推论蕴涵了当 $\| \prec \xi, \eta \succ \| < \delta$ 时

$$\|F(\alpha+\xi, \beta+\eta) - F(\alpha, \beta+\eta) - dF^1_{\prec\alpha,\beta\succ}(\xi)\| \leqslant \varepsilon\|\xi\|.$$

对

$$H(\eta) = F(\alpha, \beta+\eta) - dF^2_{\prec\alpha,\beta\succ}(\eta)$$

的相似推理使我们得到, 当 $\| \prec 0, \eta \succ \| < \delta$ 时

$$\|F(\alpha, \beta+\eta) - F(\alpha, \beta) - dF^2_{\prec\alpha,\beta\succ}(\eta)\| \leqslant \varepsilon\|\eta\|.$$

结合此两个不等式, 我们有当 $\| \prec \xi, \eta \succ \| < \delta$ 时

$$\|\Delta F_{\prec\alpha,\beta\succ}(\xi,\eta) - T(\prec \xi, \eta \succ)\| \leqslant \varepsilon\| \prec \xi, \eta \succ \|,$$

其中, $T = dF^1_{\prec\alpha,\beta\succ} \circ \pi_1 + dF^2_{\prec\alpha,\beta\succ} \circ \pi_2$. 这就是说, $\Delta F_{\prec\alpha,\beta\succ} - T = \mathfrak{o}$, 故 $dF_{\prec\alpha,\beta\succ}$ 存在并等于 T. □

对多于两个因子空间的这种定理是一个推论.

定理 8.3 如果 A 是 $\prod_1^n V_i$ 的一个开子集, 而 $F: A \to W$ 对每个 $i = 1, \cdots, n$ 及每个 $\boldsymbol{\alpha} \in A$ 其偏微分 $dF^i_{\boldsymbol{\alpha}}$ 存在, 并作为 $\boldsymbol{\alpha} = \prec \alpha_1, \cdots, \alpha_n \succ$ 的函数是连续的. 于是 $dF_{\boldsymbol{\alpha}}$ 存在并在 A 上连续. 如果 $\boldsymbol{\xi} = \prec \xi_1, \cdots, \xi_n \succ$, 则 $dF_{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \sum_1^n dF^i_{\boldsymbol{\alpha}}(\xi_i)$.

证明 由定理 8.2, dF_α^1 和 dF_α^2 的存在性与连续性表明 $dF_\alpha^1 \circ \pi_1 + dF_\alpha^2 \circ \pi_2$ 是将 F 看成是头两个变量而其他变量固定不动时的函数时的微分. 由于它是连续函数的和, 故它自身对 α 连续, 从而我们现在可再次应用定理把 dF_α^3 加进这个偏微分的和中, 最后得到 $\sum_1^3 dF_\alpha^i \circ \pi_i$ 是 F 在三因子空间 $V_1 \times V_2 \times V_3$ 上的偏微分, 等等 (这是归纳法的口头用语). □

作为这两个定理的运用示例, 我们将推导一般乘积规则 (虽然基于 Δ 估值方法的直接证明是安全可行的). 一般乘积不过是一个有界双线性映射 $\omega : X \times Y \to W$, 其中 X, Y 和 W 都是赋范线性空间. 这里的有界性不等式为 $\|\omega(\xi,\eta)\| \leqslant b\|\xi\|\|\eta\|$.

我们首先证明 ω 是可微的.

引理 8.3 *一个有界双线性映射 $\omega : X \times Y \to W$ 是处处可微的, 并且*

$$d\omega_{\prec \alpha,\beta \succ}(\xi,\eta) = \omega(\alpha,\eta) + \omega(\xi,\beta).$$

证明 固定 β 不动, 则 $g_\beta(\xi) = \omega(\xi,\beta)$ 在 $\mathrm{Hom}(X,W)$ 中, 从而处处可微, 并等于它自己的微分. 这就是说, $d\omega^1$ 处处存在且 $d\omega^1_{\prec \alpha,\beta \succ}(\xi) = \omega(\xi,\beta)$. 由于 $\beta \mapsto g_\beta$ 是个有界线性映射, 故 $d\omega^1_{\prec \alpha,\beta \succ} = g_\beta$ 是 $\prec \alpha,\beta \succ$ 的连续函数, 类似地, $d\omega^2_{\prec \alpha,\beta \succ}(\eta) = \omega(\alpha,\eta)$ 且 $d\omega^2$ 连续. 现在此引理便是定理 8.2 的直接推论了. □

如果把 $\omega(\xi,\eta)$ 想成是 ξ 和 η 的乘积, 则两个函数 $g(\zeta)$ 和 $h(\zeta)$ 的乘积便是 $\omega(g(\zeta),h(\zeta))$, 其中 g 从赋范线性空间 V 中开子集 A 到 X 而 h 从 A 到 Y. 乘积规则现在正是我们预想的那样: 积的微分是第一个乘以第二个的微分加上第二个乘以第一个的微分.

定理 8.4 *如果 $g : A \to X$ 和 $h : A \to Y$ 在 β 可微, 则乘积 $F(\zeta) = \omega(g(\zeta),h(\zeta))$ 也如此, 并且*

$$dF_\beta(\zeta) = \omega(g(\beta), dh_\beta(\zeta)) + \omega(dg_\beta(\zeta), h(\beta)).$$

证明 这是定理 8.1, 引理 8.3 和链规则的直接推论. □

习题

8.1 求弧 $\prec \sin t, \cos t, t^2 \succ$ 在 $t = 0; t = \pi/2$ 的切向量. 上述参数弧在所给两个点的微分是什么? 即, 如果 $f(t) = \prec \sin t, \cos t, t^2 \succ$, df_0 和 $df_{\pi/2}$ 是什么?

8.2 给出引理 8.1 的详细证明.

8.3 公式

$$dF_\alpha(\xi)=\sum_1^n dF_\alpha^i(\xi)$$

由于有恒等式

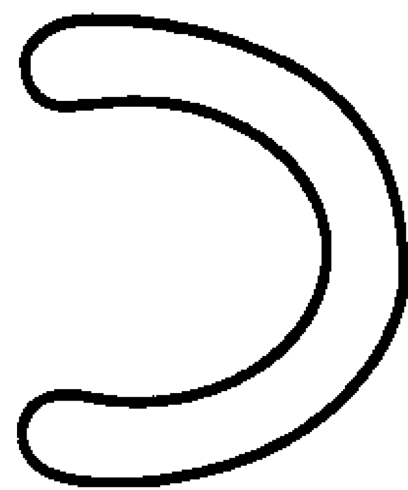
$$\xi=\sum_1^n \theta_i(\xi_i)$$

以及偏导数的定义大体上是显然的, 但不管怎样, 请写出一个明确而详细的证明.

8.4 设 F 为从 n 维向量空间 V 到一个有限维向量空间 W 的可微映射, 定义 $G:V\times W\to W$ 为 $G(\xi,\eta)=\eta-F(\xi)$. 因而 F 在 $V\times W$ 中的图像便是 G 的零集合. 证明 $dG_{\prec\alpha,\beta\succ}$ 的零空间对每个 $\prec\alpha,\beta\succ\in V\times W$ 具有维数 n.

8.5 设 $F(\xi,\eta)$ 为定义在乘积 $A\times B$ 上的连续可微函数, 其中 B 是一个球而 A 是个开集. 假定对所有 $\prec\alpha,\beta\succ\in A\times B$ 有 $dF^2_{\prec\alpha,\beta\succ}=0$. 证明 F 与 η 无关. 这就是说, 有一个定义在 A 上的连续可微函数 $G(\xi)$ 使得 $F(\xi,\eta)$ 在 $A\times B$ 上等于 $G(\xi)$.

8.6 考虑 $\mathbb{R}^2$ 中一个如右端指出的区域以证明在 $\mathbb{R}^2$ 的一个开集 A 上存在一个函数 $f(x,y)$ 使得处处有



$$\frac{\partial f}{\partial y}=0$$

使得 $f(x,y)$不仅仅只是 x 的函数.

8.7 设 $f(\xi,\eta,\zeta)$ 为三个向量变元的函数; 对于固定的 γ 令 $G(\xi,\eta)=F(\xi,\eta,\gamma)$. 证明偏微分 $dF^1_{\prec\alpha,\beta,\gamma\succ}$ 存在当且仅当 $dG^1_{\prec\alpha,\beta\succ}$ 存在, 且此时它们相等.

8.8 给出定理 8.3 的更为细致的证明. 就是说, 叙述归纳假设并证明此定理由它和定理 8.2 得出. 如果你在你的推理中是一丝不苟的, 你应需要上面习题那样的形式.

8.9 设 f 是从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的可微映射. 将 $\mathbb{R}^2$ 看作 $\mathbb{R}\times\mathbb{R}$ 时, 证明它的两个偏微分不过是用它的偏导数作乘法. 推广到 n 维情形. 证明对从 $\mathbb{R}^2$ 到一个一般的向量空间 V 的映射 F 上述结果仍然是对的, 但这时的偏导数为向量.

8.10 给出定理 8.4 证明的细节.

3.9 微分和 $\mathbb{R}^n$

我们现在把前面最后两节的结果用于涉及笛卡儿空间 $\mathbb{R}^n$ 的映射上; 笛卡儿空间是有限维理论的常用对象. 我们从定义域开始讨论.

定理 9.1 如果 F 是从 $\mathbb{R}^n$(的一个开子集) 到一个赋范线性空间 W 的映射, 则 F 沿第 j 个标准基向量 δ^j 方向的方向导数恰好是偏导数 $\partial F/\partial x_j$, 而第 j 个偏微分是以 $\partial F/\partial x_j$ 作的乘法:$dF_{\boldsymbol{a}}^j(h) = h(\partial F/\partial x_j)(\boldsymbol{a})$. 更准确地, 如果上面三个对象中的任一个在 $\boldsymbol{a}$ 存在则它们全都存在, 并满足上面的关系.

证明 我们有

$$\frac{\partial F}{\partial x_j}(\boldsymbol{a}) = \lim_{t\to 0}\frac{F(a_1,\cdots,a_{j+t},\cdots,a_n) - F(a_1,\cdots,a_j,\cdots,a_n)}{t}$$
$$= \lim_{t\to 0}\frac{F(\boldsymbol{a}+t\delta^j) - F(\boldsymbol{a})}{t} = D_{\delta j}F(\boldsymbol{a}).$$

另外, 由于 F 在 $\boldsymbol{a}+\mathbb{R}\delta^j$ 的限制是条参数弧, 其在 0 的微分由定义为 F 在 $\boldsymbol{a}$ 的第 j 个偏微分, 并且其在 0 的切向量为我们已经计算过的 $(\partial F/\partial x_j)(\boldsymbol{a})$, 本定理的其余部分由定理 7.1 推出. □

综合这个定理和定理 7.2, 我们得到下列结果.

定理 9.2 如果 $V = \mathbb{R}$ 和 F 在 $\boldsymbol{a}$ 为可微的, 则偏导数 $(\partial F/\partial x_j)(\boldsymbol{a})$ 都存在而且在 $\boldsymbol{a}$ 偏导数的 n 联组 $\{\partial F/\partial x_j(\boldsymbol{a})\}_n^1$ 为 dF_a 的骨架. 特别

$$D_yF(\boldsymbol{a}) = \sum_{j=1}^{n} y_i\frac{\partial F}{\partial x_j}(\boldsymbol{a}).$$

证明 $dF_{\boldsymbol{a}}(\delta^i) = D_{\delta^i}F(\boldsymbol{a}) = (\partial F/\partial x_i)(\boldsymbol{a})$, 如我们在上面指明的, 我们有

$$D_{\boldsymbol{y}}F(\boldsymbol{a}) = dF_{\boldsymbol{a}}(\boldsymbol{y}) = dF_{\boldsymbol{a}}\left(\sum_1^n y_i\delta^i\right) = \sum_1^n y_i dF_{\boldsymbol{a}}(\delta^i) = \sum_1^n y_i\frac{\partial F}{\partial x_i}(\boldsymbol{a}).$$

在此我们已经做了全部显示的 $dF_{\boldsymbol{a}}$ 是由它的骨架 $\{dF_{\boldsymbol{a}}(\delta^i)\}$ 定义的线性组合映射 (见第一章定理 1.2), 其中 $T(\delta^i) = dF_{\boldsymbol{a}}(\delta^i)$ 现在被认为偏导数 $(\partial F/\partial x_i)(\boldsymbol{a})$.□

上面公式表明偏导数记号的不规范性: 注意一下如果我们试图计算 $dF_{\alpha}(\boldsymbol{x})$ 时它是如何暴露出来的. 记号 $D_{\delta j}F$ 准确但繁琐. 其他的记号为 F_j 和 D_jF. 每个都有它的问题, 但第二个记号大体是难点最少的. 使用它后我们的公式便读成 $dF_{\boldsymbol{a}}(\boldsymbol{y}) = \sum_{j=1}^n y_jD_jF(\boldsymbol{a})$.

在相反的方向方面我们有定理 8.3 的相应的特定描述.

定理 9.3 如果 A 是 $\mathbb{R}^n$ 的一个开子集, F 是从 A 到赋范线性空间 W 的一个映射, 使得所有的偏导数 $(\partial F/\partial x_j)(\boldsymbol{a})$. 存在并在 A 上连续, 则 F 在 A 上连续可微.

证明 由于 F 的第 j 个偏微分是简单地乘以 $\partial F/\partial x_j$, 我们 (由定理 9.1) 有了所有偏微分 $dF_{\boldsymbol{a}}^j$ 在 A 的存在性和连续性. 因而定理 9.3 成了定理 8.3 的一个特殊情形. □

现在假如 F 的值域空间也是一个笛卡儿空间, 故 F 是从 $\mathbb{R}^n$ 的一个开子集 A 到 $\mathbb{R}^n$ 的映射. 那么 $dF_{\boldsymbol{a}}$ 在 $\mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ 中. 为了计算的目的我们需要用它们的矩阵来表示从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性映射, 因而求微分 $T = dF_{\boldsymbol{a}}$ 的矩阵 $\boldsymbol{t}$ 具有极其重要性. 称这个矩阵为 F 在 $\boldsymbol{a}$ 的雅可比矩阵.

$\boldsymbol{t}$ 的列构成 $dF_{\boldsymbol{a}}$ 的骨架, 我们在上面已看到过此骨架是偏导数 $(\partial F/\partial x_j)(\boldsymbol{a})$ 的 n 联组. 如果我们不严格地写 m 联组值 F 为一个 m 联组的函数, $F = \prec f_1, \cdots, f_m \succ$, 则按照引理 8.1, $\boldsymbol{t}$ 的第 j 列是 m 联组

$$\frac{\partial F}{\partial x_j}(\boldsymbol{a}) = \prec \frac{\partial f_1}{\partial x_j}(\boldsymbol{a}), \cdots, \frac{\partial f_m}{\partial x_j}(\boldsymbol{a}) \succ .$$

因此,

定理 9.4 设 F 是从 $\mathbb{R}^n$ 中一个开子集到 $\mathbb{R}^m$ 的映射, 且假设 F 在 $\boldsymbol{a}$ 可微. 于是 $dF_{\boldsymbol{a}}$ 的矩阵 (F 在 $\boldsymbol{a}$ 的雅可比矩阵) 由

$$t_{ij} = \frac{\partial f_i}{\partial x_j}(\boldsymbol{a})$$

给出. 如果我们用记号 $y_i = f_i(\boldsymbol{x})$, 我们有

$$t_{ij} = \frac{\partial y_i}{\partial x_j}(\boldsymbol{a}).$$

如果我们也有从 $\mathbb{R}^m$ 中一个开集 B 到 $\mathbb{R}^l$ 的一个可微映射 $z = G(\boldsymbol{y}) = \prec g_1(y), \cdots, g_l(y) \succ$, 则相似地, $dG_{\boldsymbol{b}}$ 有矩阵

$$\frac{\partial g_k}{\partial y_i}(\boldsymbol{b}) = \frac{\partial z_k}{\partial y_i}(\boldsymbol{b}).$$

又, 如果 B 包含了 $\boldsymbol{b} = F(\boldsymbol{a})$, 则复合函数规则

$$d(G \circ F)_{\boldsymbol{a}} = dG_{\boldsymbol{b}} \circ dF_{\boldsymbol{a}}$$

具有矩阵形式

$$\frac{\partial z_k}{\partial x_j}(\boldsymbol{a}) = \sum_{i=1}^{m} \frac{\partial z_k}{\partial y_i}(\boldsymbol{b}) \frac{\partial y_i}{\partial x_j}(\boldsymbol{a}),$$

或简单地

$$\frac{\partial z_k}{\partial x_j} = \sum_{i=1}^{m} \frac{\partial z_k}{\partial y_i} \frac{\partial y_i}{\partial x_j}.$$

这是微积分中链规则的通常形式. 我们看出这只是将线性映射的复合表达为矩阵的相乘.

在第 8 节中我们曾看到单实变函数 f 的普通导数 $f'(a)$ 是微分 df_a 的骨架, 从而把这个关系推广并定义一个 n 个实变量的函数 F 的导数 $F'(\boldsymbol{a})$ 为 $dF_{\boldsymbol{a}}$ 的骨架是有道理的, 故而 $F'(\boldsymbol{a})$ 像我们上面曾看到的那样为偏导数的 n 联组 $\{(\partial F/\partial x_i)(\boldsymbol{a})\}_1^n$. 特别, 如果 F 是从 $\mathbb{R}^n$ 中一个开子集到 $\mathbb{R}^m$, 则 $F'(\boldsymbol{a})$ 是 F 在 $\boldsymbol{a}$ 的雅可比矩阵. 这给了矩阵链规则的标准形式

$$(G \circ F)'(\boldsymbol{a}) = G'(F(\boldsymbol{a}))F'(\boldsymbol{a}).$$

一些作者对我们叫做微分的东西使用 '导数' 这个词, 但是这是对单变情形中的传统意义的一个变化, 从而我们宁愿保持上面讨论中的区别: 微分 dF_{α} 是近似于 ΔF_{α} 的线性映射, 并且当定义域和值域空间为笛卡儿时导数 $F'(\boldsymbol{a})$ 必定是此线性映射的矩阵. 然而, 我们将继续采用雅可比术语.

假设现在 A 是有限维向量空间 V 中的一个开子集, 而 $H: A \to W$ 在 $\alpha \in A$ 可微. 假如 W 也是有限维的, 并且 $\varphi: V \to \mathbb{R}^n$ 和 $\psi: W \to \mathbb{R}^m$ 是任一坐标同构. 如果 $\overline{A} = \varphi[A]$, 则 $\overline{A}$ 是 $\mathbb{R}^n$ 中一个开子集, 而 $\overline{H} = \psi \circ H \circ \varphi^{-1}$ 是从 $\overline{A}$ 到 $\mathbb{R}^m$ 的一个映射, 它在 $\boldsymbol{a} = \varphi(\alpha)$ 可微, 微分 $d\overline{H}_{\boldsymbol{a}} = \psi \circ dH_{\alpha} \circ \varphi^{-1}$. 于是 $d\overline{H}_{\boldsymbol{a}}$ 由它的雅可比矩阵 $\{(\partial h^i/\partial x_j)(\boldsymbol{a})\}$ 给出, 我们称它为 H 关于 V 和 W 中所选基的雅可比矩阵. V 和 W 中的基变化将按 2.4 节中给出的规则改变这个雅可比矩阵.

如果 F 是 $\mathbb{R}^n$ 到其自身的映射, 则雅可比矩阵 $(\partial f^i/\partial x_j)(\boldsymbol{a})$ 的行列式被叫做 F 在 $\boldsymbol{a}$ 的雅可比. 它被表记为

$$\frac{\partial(f^1, \cdots, f^n)}{\partial(x_1, \cdots, x_n)}(\boldsymbol{a}) \quad \text{或者} \quad \frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}(\boldsymbol{a}),$$

这里暗含有 $y_i = f^i(\boldsymbol{x})$. 另一个记号是 $J_F(\boldsymbol{a})$(或简单地 $J(\boldsymbol{a})$; 这时 F 是不言自明的). 但是有时它被用来表示微分 $dF_{\boldsymbol{a}}$, 此时我们转而使用 $\det J_F(\boldsymbol{a})$.

如果 $F(x) = \prec x_1^2 - x_2^2, 2x_1x_2 \succ$, 那么它的雅可比矩阵是

$$\begin{bmatrix} 2x_1 & -2x_2 \\ 2x_2 & x_1 \end{bmatrix},$$

而 $\det J(\boldsymbol{x}) = 4(x_1^2 + x_2^2) = 4(\|\boldsymbol{x}\|_2)^2$.

习题

9.1 用参数弧概念的类比, 我们定义在赋范线性空间 W 中的一个光滑的参数 (化的) 二维曲面为一个从矩形 $I \times J \in \mathbb{R}^2$ 到 W 的连续可微映射 Γ. 假如 $I \times J = [-1,1] \times [-1,1]$, 构造 Γ 在 W 中的值域在点 $\Gamma(0,0)$ 处切空间的定义. 证明两个向量

$$\frac{\partial \Gamma}{\partial x}(0,0) \quad 和 \quad \frac{\partial \Gamma}{\partial y}(0,0)$$

是这个切空间的基.(不要把这个当作你的定义.)

9.2 把上一个习题推广到赋范线性空间 W 中的光滑参数 n 维曲面.

9.3 计算映射 $\prec x, y \succ \mapsto \prec x^2, y^2, (x+y)^2 \succ$ 的雅可比矩阵. 证明除原点外其秩为 2.

9.4 设 $F = \prec f^1, f^2, f^3 \succ$ 为从 $\mathbb{R}^3$ 到 $\mathbb{R}^3$, 其定义为

$$f^1(x,y,z) = x + y + z, \quad f^2(x,y,z) = x^2 + y^2 + z^2$$

和

$$f^3(x,y,z) = x^3 + y^3 + z^3.$$

计算 F 在 $\prec a, b, c \succ$ 的雅可比. 证明除了三个坐标中有两个相等外它是非异的. 描述奇点的轨迹.

9.5 计算从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的映射 $F : \prec x, y \succ \mapsto \prec (x+y)^2, y^2 \succ$ 在点 $\prec 1, -1 \succ$; $\prec 1, 0 \succ$; $\prec a, b \succ$ 的雅可比. 计算 $G : \prec s, t \succ \mapsto \prec s - t, s + t \succ$ 在 $\prec u, v \succ$ 的雅可比.

9.6 在上面的习题中计算复合 $F \circ G$ 和 $G \circ F$. 计算 $F \circ G$ 在 $\prec y, v \succ$ 的雅可比. 计算 F 和 G 的雅可比的相应乘积.

9.7 计算由 $x = r\cos\theta, y = r\sin\theta, z = z$ 定义的映射 T 的雅可比矩阵和行列式. 对此映射复合上一个函数 $f(x,y,z)$ 便给出了一个新的函数:

$$g(r,\theta,z) = f(r\cos\theta, r\sin\theta, z).$$

就是说, $g = f \circ T$. 这个复合(替换) 称为换成 $\mathbb{R}^3$ 中的柱坐标.

9.8 计算极坐标变换 $\prec r, \theta \succ \mapsto \prec x, y \succ, x = r\cos\theta, y = r\sin\theta$ 的雅可比行列式.

9.9 变换为球面坐标是由 $x=r\sin\varphi\cos\theta, y=r\sin\varphi\sin\theta, z=r\cos\theta$ 给出. 计算雅可比

$$\frac{\partial(x,y,z)}{\partial(r,\varphi,\theta)}.$$

9.10 写出对下面特殊情形的链规则:

$$\frac{dw}{dt}=?,\quad 其中w=F(x,y),x=g(t),y=h(t).$$

求当 $w=F(x_1,\cdots,x_n)$ 而 $x_i=g_i(t),i=1,\cdots,n$ 时的 dw/dt. 求当 $w=F(x,y),x=g(u,v),y=h(u,v)$ 时的 $\partial w/\partial u$. 在 $g(u,v)=u$ 的特殊情形可以重写为

$$\frac{\partial}{\partial x}F(x,h(x,v)).$$

计算它.

9.11 如果 $w=f(x,y),x=r\cos\theta,y=r\sin\theta$, 证明

$$\left[r\frac{\partial w}{\partial r}\right]^2+\left[\frac{\partial w}{\partial\theta}\right]^2=\left[\frac{\partial w}{\partial x}\right]^2+\left[\frac{\partial w}{\partial y}\right]^2.$$

3.10 初步应用

由标准微积分中的初等极大极小理论所作的推广少有变化, 此刻我们只做简短的讨论.

定理 10.1 *设 F 为赋范线性空间 V 中开子集 A 上的一个实函数, 并假定 F 在 A 中点 α 取得一相对极大值, 并且 dF_α 存在. 于是 $dF_\alpha=0$.*

证明 由定义 $D_\xi F(\alpha)$ 为函数 $\gamma(t)=F(a+t\xi)$ 的导数 $\gamma'(0)$, 而 γ 的定义域是 0 在 $\mathbb{R}$ 中的一个邻域. 由于 $\gamma(t)$ 在 0 有一个相对极大值, 我们由初等微积分有 $\gamma'(0)=0$. 因此 $dF_\alpha(\xi)=D_\xi F(\alpha)=0$ 对所有 ξ 成立, 故 $dF_\alpha=0$. □

使得 $dF_\alpha=0$ 的点 α 被称作*判别点*.此定理说, 一个可微实函数只有在一个判别点才能有一个内部极值.

如果 V 是 $\mathbb{R}^n$, 则上述论证表明一个实函数 F 仅当在 $\boldsymbol{a}$ 的偏导数 $(\partial F/\partial x_i)(\boldsymbol{a})$ 全都为 0 时它才能在 $\boldsymbol{a}$ 有一个相对极大 (或极小), 因而像在初等微积分中那样, 这常常提供了计算极大 (或极小) 值的一个办法. 例如, 假如我们要证明从对固定体积 V 使表面最小的观点来说, 立方体是长方体中最有效的一个. 如

果各边长为 x, y, z, 我们有 $V = xyz$ 和 $A = 2(xy + xz + yz) = 2(xy + V/y + V/x)$. 于是由 $0 = \partial A/\partial x = 2(y - V/x^2)$ 我们有 $V = yx^2$, 相似地, $\partial A/\partial y = 0$ 表明 $V = xy^2$. 因此 $yx^2 = xy^2$, 且由于 x 和 y 都不能为 0, 从而有 $x = y$. 于是 $V = yx^2 = x^3, x = V^{1/3} = y$. 最后在 $V = xyz$ 中替换表明 $z = V^{1/3}$. 我们的判别构形因而是个立方体, 其极小面积 $A = 6V^{2/3}$.

上面已假定了 A 在某点 $\prec x, y, z \succ$有了一个绝对极小值. 读者可以证明如果 x, y, z 任一趋向 0 或 ∞ 时 $A \to \infty$ 意味着此极小值确实存在, 并在证明中享受乐趣.

我们在 12,15 和 16 节要回到这个决定判别点的问题.

条件 $dF_\alpha = 0$ 对一个内部极大或极小是必要而非充分的. 读者应想起在初等微积分中的一个充分条件: 如果 $f'(x) = 0$ 且 $f''(x) < 0 (> 0)$ 则 x 是 f 的相对极大 (极小) 点. 我们在第 16 节将证明其相应的一般性定理. 但现在有更多的可能性; 在其中有类比的充分条件, 即如果 $dF_\alpha = 0$ 且 d^2F_α 负 (正) 定 (这里把 d^2F_α 当作 V 上的一个二次型), 则 α 是 F 的一个相对极大 (极小) 点.

下面我们考虑图像的切平面概念. 曲线的切线或曲面的切平面的计算通常被当作导数的几何应用, 从而我们也把这个作为在这里考虑这种一般性问题的十分正当的理由.

设 F 是从赋范线性空间 V 的一个开子集 A 到赋范线性空间 W 的一个映射. 当我们将 F 看作 $V \times W$ 上的一个图像时我们把它想成是在定义域 A "上方" 的一个 "曲面"S, 这推广了两个实变量的实函数在 $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R}$ 中图像的几何解释. 投影 $\pi_1 : V \times W \to V$ 投 S"下" 到 A 上,

$$\prec \xi, F(\xi) \succ \underset{\pi_1}{\longmapsto} \xi,$$

而映射 $\xi \mapsto \prec \xi, F(\xi) \succ$ 给出了 S 中在 ξ"上方" 的点. 我们的几何想像把 V 看作是 $V \times W$ 中的平面 (子空间)$V \times \{0\}$, 正如同我们习惯地把 $\mathbb{R}$ 形象地当作 $\mathbb{R}^2$ 中的实轴 $\mathbb{R} \times \{0\}$.

现在我们假定 F 在 α 可微. 在我们第 6 节中的预备性讨论中需提到线性函数 dF_α(的图像) 是函数 ΔF_α 在 $V \times W$ 中的 (图像的) 切平面, 而且它的经过 $\prec \alpha, F(\alpha) \succ$ 的平移 M 是曲面 S 在 $\prec \alpha, F(\alpha) \succ$ 的切平面方程, 这里的 S 就是 F(的图像). 此平面的方程是 $\eta - F(\alpha) = dF_\alpha(\xi - \alpha)$, 因而是仿射函数 $G(\xi) = dF_\alpha(\xi - \alpha) + F(\alpha)$(的图像). 我们知道 dF_α 是 $\mathrm{Hom}(V, W)$ 中惟一的 T 使得 $\Delta F_\alpha(\zeta) = T(\zeta) + \mathfrak{o}(\zeta)$, 从而如果我们令 $\zeta = \xi - \alpha$, 则容易看出这等同于说 G

是惟一的从 V 到 W 的仿射映射使得

$$F(\xi)-G(\xi)=\mathfrak{o}(\xi-\alpha).$$

这就是说 M 是 V 上在 $\prec\alpha,F(\alpha)\succ$ 周围, 以 $\mathfrak{o}$ - 近似的意义 "贴合" 于曲面 S 的惟一平面.

然而有更进一步的几何事实大大地增强了我们的感觉: 这真正是切平面.

定理 10.2 具方程 $\eta-F(\alpha)=dF_\alpha(\xi-\alpha)$ 的平面恰好是 $V\times W$ 中通过 $\prec\alpha,F(\alpha)\succ$ 的所有那些直线的并, 这些直线切于曲面 $S=F$ 的图像上通过此点的光滑曲线. 换句话说, $V\times W$ 的子空间 dF_α 的向量正好是在 S 中并通过 $\prec\alpha,F(\alpha)\succ$ 的曲线的切向量.

证明 这几乎是显而易见的. 如果 $\prec\xi,\eta\succ\in dF_\alpha$, 则弧

$$\gamma(t)=\prec\alpha+t\xi,F(\alpha+t\xi)\succ$$

在 S 中并位于 V 中直线 $t\mapsto\alpha+t\xi$ 之上, 它以 $\prec\xi,dF_\alpha(\xi)\succ=\prec\xi,\eta\succ$ 为其在 $\prec\alpha,F(\alpha)\succ$ 的切向量; 这些可由引理 8.1 和定理 8.2 推导出.

反之, 如果 $t\mapsto\prec\lambda(t),F(\lambda(t))\succ$ 为 S 中通过 α 的任一条光滑曲线, 具有 $\lambda(t_0)=\alpha$, 则它在 $\prec\alpha,F(\alpha)\succ$ 的切向量为

$$\prec\lambda'(t_0),dF_\alpha(\lambda'(t_0))\succ,$$

这是 dF_α(的图像) 中的向量. □

作为上面所讨论的一般切平面的例子, 设 $F=\prec f_1,f_2\succ$ 为 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的映射, 其定义为 $f_1(x)=(x_1^2-x_2^2)/2, f_2(x)=x_1x_2$. F 的图像是 $\mathbb{R}^4=\mathbb{R}^2\times\mathbb{R}^2$ 中在 $\mathbb{R}^2$ 上的一个曲面. 根据我们上面的讨论, 在 $\prec\boldsymbol{a},F(\boldsymbol{a})\succ$ 的切平面有方程 $y=dF_{\boldsymbol{a}}(\boldsymbol{x}-\boldsymbol{a})+F(\boldsymbol{a})$. 在 $\boldsymbol{a}=\prec 1,2\succ, dF_{\boldsymbol{a}}$ 的雅可比矩阵是

$$\begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix}_{\prec 1,2\succ}=\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix},$$

而 $F(\boldsymbol{a})=\prec -\frac{3}{2},2\succ$. 在 $\prec 1,2\succ$ 的切平面 M 的方程因而为

$$\prec y_1,y_2\succ=\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}\prec x_1-1,x_2-2\succ+\prec -\frac{3}{2},2\succ.$$

计算矩阵积, 我们得到纯量方程

$$y_1 = x_1 - 2x_2 + \left(-1 + 4 - \frac{3}{2}\right) = x_1 - 2x_2 + \frac{3}{2},$$

$$y_2 = 2x_1 + x_2 + (-2 - 2 + 2) = 2x_1 + x_2 - 2.$$

请注意, 这两个方程把仿射空间 M 表示为 $\mathbb{R}^4$ 中所有满足

$$x_1 - 2x_2 - y_1 = -\frac{3}{2}$$

的 $\prec x_1, x_2, y_1, y_2 \succ$ 组成的超平面与具有方程

$$2x_1 + x_2 - y_2 = 2$$

的超平面的交.

习题

10.1 求 $f(x,y,z) = x + y + z$ 在椭球

$$x^2 + 2y^2 + 3z^2 = 1$$

上的极大值.

10.2 求线性泛函 $f(\boldsymbol{x}) = \sum_1^n c_i x_i$ 在单位球 $\sum_1^n x_i^2 = 1$ 上的极大值.

10.3 求在 $\mathbb{R}^3$ 中两条直线

$$\boldsymbol{x} = t\delta^1 \quad \text{和} \quad \boldsymbol{y} = s \prec 1,1,1 \succ + \prec 1,0,-1 \succ$$

间的 (最小) 距离.

10.4 证明除了对某个 k 有 $b = ka$ 外在 $\mathbb{R}^n$ 中两条直线 $\boldsymbol{x} = t\boldsymbol{a} + \boldsymbol{l}$ 和 $\boldsymbol{y} = s\boldsymbol{b} + \boldsymbol{m}$ 上存在一对靠得最近的点. 我们假定 $\boldsymbol{a} \neq 0 \neq \boldsymbol{b}$. 回想如果 $\boldsymbol{b}$ 不是 $k\boldsymbol{a}$ 的形式时则根据施瓦兹不等式有 $|(\boldsymbol{a},\boldsymbol{b})| < \|\boldsymbol{a}\|_2 \|\boldsymbol{b}\|_2$.

10.5 证明原点是 $f(x,y,z) = xy + yz + zx$ 的惟一的判别点. 求通过原点的一条直线使得沿着它 0 是 f 的极大点, 再求另一条直线使沿着它 0 是 f 的一个极小点.

10.6 正文中讨论过的具给定体积 V 的长方体的面积极小化问题里, 假定了

$$A = 2\left(xy + \frac{V}{x} + \frac{V}{y}\right)$$

在第一象限的内点上有一个绝对极小值. 证明这点. 先证明如果 $\prec x, y \succ$ 以任何方式趋向边界时 $A \to \infty$:

$$x \to 0, \quad x \to \infty, \quad y \to 0 \ \text{或} \ y \to \infty.$$

10.7 设 $F:\mathbb{R}^2\to\mathbb{R}^2$ 为由

$$y_1=\sin(x_1+x_2), y_2=\cos(x_1-x_2)$$

定义的映射. 求 $\mathbb{R}^4$ 中 F 的图像在点 $\boldsymbol{a}=\prec\pi/4,\pi/4\succ$ 上的切平面的方程.

10.8 定义 $F:\mathbb{R}^2\to\mathbb{R}^2$ 为

$$y_1=\sum_1^3 x_i^2, y_2=\sum_1^3 x_i^3.$$

求 F 在 $\mathbb{R}^5$ 中图像在 $\boldsymbol{a}=\prec 1,2,-1\succ$ 上的切平面的方程.

10.9 设 $\omega(\xi,\eta)$ 为从乘积赋范线性空间 $V\times W$ 到赋范线性空间 X 的有界双线性映射. 证明 ω 在 $V\times W\times X$ 的图像 S 在点 $\prec\alpha,\beta,\gamma\succ\in S$ 的切平面方程为

$$\zeta=\omega(\xi,\beta)+\omega(\alpha,\eta)+\omega(\alpha,\beta).$$

10.10 设 F 是赋范线性空间 V 上的有界线性泛函. 证明 F^3 在 $V\times\mathbb{R}$ 中图像在点 α 上方的切平面方程能够被写成形式 $y=F^2(\alpha)(3F(\xi)-2F(\alpha))$.

10.11 证明如果在正文中所给出的切平面的一般方程被用于一个 $\mathrm{Hom}(V,W)$ 中的映射 F, 则它可以化成 F 本身的方程 $[\eta=F(\xi)]$ 而不管切点在哪里.(自然如此!)

10.12 继续习题 9.1, 证明 Γ 在 W 中的值域在 $\Gamma(0)$ 的切空间是 Γ 在 $\mathbb{R}^2\times W$ 中图像在点 $\prec 0,\Gamma(0)\succ$ 的切空间在 W 中的投影. 现在定义 Γ 在 W 内的值域在 $\Gamma(0)$ 的切平面并证明相似地为 Γ 图像的切平面的投影.

10.13 设 $F:V\to W$ 在 α 可微. 证明 dF_α 的值域是 F 在 $V\times W$ 的图像在点 $\prec\alpha,F(\alpha)\succ$ 的切空间在 W 上的投影.

3.11 隐函数定理

我们在第 9 节得到的对复合映射的雅可比的公式使我们想起早先 (第 8 节) 推导出的对复合映射微分的链规则. 雅可比公式包含我们进行乘与加的数 (偏导数); 微分链规则则包含了我们进行复合与加的线性映射 (偏微分).(如果我们利用块分解则这种相似性成了完全的形式类比.) 粗略地说, 整个微分学就是按此方式进行的. 在一元微积分中微分是从一维空间 $\mathbb{R}$ 到自身的线性映射, 因而是用一个数的乘法, 此数即导数. 在多元微积分中当我们作关于一维子空间的分解时, 我们得到这些数的分块, 即雅可比矩阵. 当我们把整个理论推广到非一维的向量空间时, 我们得到本质上一样的公式但只是用线性映射 (微分) 替换了数, 复合替换了乘法.

因一个反函数的导数是此函数导数的*互反*(倒数): 如果 $g = f^{-1}, b = f(a)$, 则 $g'(b) = 1/f'(a)$. 一个逆映射的微分是此映射微分的*复合逆*: 如果 $G = F^{-1}$ 且 $F(\alpha) = \beta$, 则 $dG_\beta = (dF_\alpha)^{-1}$.

如果方程 $g(x, y) = 0$ 隐式地定义了 y 为 x 函数 $y = f(x)$, 在初等微积分中我们学会了计算 $f'(a)$, 方法是微分

$$g(x, f(x)) \equiv 0$$

并得到

$$\frac{\partial g}{\partial x}(a, b) + \frac{\partial g}{\partial y}(a, b) f'(a) = 0,$$

其中 $b = f(a)$. 因此

$$f'(a) = -\frac{\partial g/\partial x}{\partial g/\partial y}.$$

我们将在后面看到, 如果 $G(\xi, \eta)$ 定义 η 为 ξ 的函数 $\eta = F(\xi)$, 并设 $\beta = F(\alpha)$, 则我们可以通过微分恒等式 $G(\xi, F(\xi)) = 0$ 计算*微分* dF_α, 并得一个形式上等同于上面的公式.

最后, 以完全相同的方式, 在初等微积分中解极大极小问题的所谓辅助变量法以相同的形式结构作为我们以后用拉格朗日乘子法对 "约束" 极大值的解.

在本节我们将考虑隐式定义的函数的存在性和可微性. 假如我们已有一个两个向量变元的 (向量值) 函数 $G(\xi, \eta)$, 我们需要知道是否令 $G = 0$ 便定义了 η 为 ξ 的一个函数, 即是否存在一个惟一的函数 F 使得 $G(\xi, F(\xi))$ 恒等为 0. 假如这样一个 "隐式定义的" 函数 F 存在且处处均可微, 我们便能对方程 $G(\xi, F(\xi)) = 0$ 或 $G \circ \prec I, F \succ = 0$ 的微分来计算 F 在 α 的微分. 我们得到 $dG^1_{\prec\alpha,\beta\succ} \circ dI_\alpha + dG^2_{\prec\alpha,\beta\succ} \circ dF_\alpha = 0$, 其中我们已令 $\beta = F(\alpha)$. 如果 dG^2 可逆, 我们解出 dF_α, 得到

$$dF_\alpha = -(dG^2_{\prec\alpha,\beta\succ})^{-1} \circ dG^1_{\prec\alpha,\beta\succ}.$$

留意一下, 这与我们上面所复习的初等微积分中相应的表达式具有相同的形式, 而且上面的计算强烈地提示, 我们将需要 $(dG^2_{\prec\alpha,\beta\succ})^{-1}$ 的存在性以作为在 $\prec \alpha, \beta \succ$ 附近惟一被定义的隐函数存在性的必要条件. 由于 β 等于 $F(\alpha)$, 我们还需要 $G(\alpha, \beta) = 0$. 这些考虑将把我们引向正确的定理, 然而其证明的部分内容不得不被推延到下一章. 我们在这里能证明的是如果存在一个隐式定义的函数则它必是可微的.

定理 11.1 设 V, W 和 X 为赋范线性空间, 并设 G 为从 $V\times W$ 中的开子集 $A\times B$ 到 X 的映射. 假如 F 为从 A 到 B 连续映射, 而它是由方程 $G(\xi,\eta)=0$ 隐式定义的, 即在 A 上满足 $G(\xi,F(\xi))=0$. 最后, 假如 G 在 $\prec\alpha,\beta\succ$ 可微, 其中 $\beta=F(\alpha)$, 且 $dG^2_{\prec\alpha,\beta\succ}$ 可逆. 于是 F 在 α 可微, 并且 $dF_\alpha=-(dG^2_{\prec\alpha,\beta\succ})^{-1}\circ dG^1_{\prec\alpha,\beta\succ}$.

证明 令 $\eta=\Delta F_\alpha(\xi)$, 故 $G(\alpha+\xi,\beta+\eta)=G(\alpha+\xi,F(\alpha+\xi))=0$. 于是

$$\begin{aligned}0&=G(\alpha+\xi,\beta+\eta)-G(\alpha,\beta)=\Delta G_{\prec\alpha,\beta\succ}(\xi,\eta)\\&=dG_{\prec\alpha,\beta\succ}(\xi,\eta)+\mathfrak{o}(\xi,\eta)\\&=dG^1_{\prec\alpha,\beta\succ}(\xi)+dG^2_{\prec\alpha,\beta\succ}(\eta)+\mathfrak{o}(\prec\xi,\eta\succ).\end{aligned}$$

将 T^{-1} 用于此方程, 其中 $T=dG^2_{\prec\alpha,\beta\succ}$, 并解出 η, 我们得到

$$\eta=-T^{-1}(dG^1_{\prec\alpha,\beta\succ}(\xi))+\mathfrak{O}(\mathfrak{o}(\prec\xi,\eta\succ)).$$

这个方程具有形式 $\eta=\mathfrak{O}(\xi)+\mathfrak{o}(\prec\xi,\eta\succ)$, 并且由于因 F 在 α 的连续性 $\eta=\Delta F_\alpha(\xi)$ 是个无穷小 $\mathfrak{I}(\xi)$, 从而引理 5.1 和 5.2 首先表明 $\eta=\mathfrak{O}(\xi)$, 然后表明 $\prec\xi,\eta\succ=\mathfrak{O}(\xi)$. 因此 $\mathfrak{O}(\mathfrak{o}(\prec\xi,\eta\succ))=\mathfrak{O}((\mathfrak{o}\mathfrak{O}(\xi)))=\mathfrak{o}(\xi)$, 从而我们有

$$\Delta F_\alpha(\xi)=\eta=S(\xi)+\mathfrak{o}(\xi),$$

其中 $S=-(dG^2_{\prec\alpha,\beta\succ})^{-1}\circ dG^1_{\prec\alpha,\beta\succ}$ 为 $\mathrm{Hom}(V,W)$ 的一个元素. 所以 F 在 α 可微且 dF_α 具有所断言的值. □

在下一章我们要证明如果 V, W 和 X 为有限维的, 并且 G 是连续可微的从 $V\times W$ 中开子集 $A\times B$ 到 X 的映射使得在点 $\prec\alpha,\beta\succ$ 有 $G(\alpha,\beta)=0$ 和 $dG^2_{\prec\alpha,\beta\succ}$ 可逆, 则存在一个从 α 的邻域 M 到 B 的惟一确定的连续映射 F, 使得在 M 上 $F(\alpha)=\beta$ 和 $G(\xi,F(\xi))=0$; 而此证明是作为不动点定理的一个应用. 对于下一章要学到的更广泛的一类空间即完备的赋范线性空间, 相同的定理也是对的. 对这些空间, 如果 T^{-1} 存在, 则 S^{-1} 也存在, 其中 S 为所有充分靠近 T 的映射, 并且 $S\mapsto S^{-1}$ 连续. 因此 $dG^2_{\prec\mu,\nu\succ}$ 对所有充分靠近 $\prec\alpha,\beta\succ$ 的 $\prec\mu,\nu\succ$ 可逆, 那么上面的定理表明 F 在 α 的一个邻域上可微. 另外, 对 $dF:\mu\mapsto dF_n$ 的定理给出的公式中只涉及了连续映射, 从而知道 F 实际上在 α 附近为连续可微. 这些结论构成了隐函数定理, 现在我们来叙述它.

定理 11.2 设 V, W 和 X 为有限维的 (或更一般为完备的) 赋范线性空间, 设 $A\times B$ 为 $V\times W$ 的一个开子集, 并设 $G: A\times B\to X$ 为连续可微. 假如在

$A\times B$ 的点 $\prec\alpha,\beta\succ$ 我们有 $G(\alpha,\beta)=0$ 并且 $dG^2_{\prec\alpha,\beta\succ}$ 可逆. 于是存在 α 的一个球 M 和从 M 到 B 的一个惟一确定的连续可微映射 F, 使得在 M 上 $F(\alpha)=\beta$ 和 $G(\xi,F(\xi))=0$.

所谓的逆映射定理是隐函数定理的一个特殊情形.

定理 11.3 设 H 为从有限维 (或完备) 赋范线性空间 W 中一个开子集 B 到线性赋范线性空间 V 的一个连续可微映射, 并假设它的微分在点 β 可逆. 于是 H 自身在 β 邻近可逆, 那就是说, 存在 $\alpha=H(\beta)$ 的一个球 M 和一个从 M 到 B 的惟一确定的连续可微函数 F 使得 $F(\alpha)=\beta$ 并且 $H(F(\xi))=\xi$ 在 M 上成立.

证明 令 $G(\xi,\eta)=\xi-H(\eta)$. 于是 G 是从 $V\times B$ 到 V 的一个连续可微映射并且 $dG^2_{\prec\alpha,\beta\succ}=-dH_\beta$ 可逆, 隐函数定理给了我们一个在 α 的球 M 和一个惟一确定的连续可微映射 $F:M\to B$ 使得 $F(\alpha)=\beta$ 和 $0=G(\xi,F(\xi))=\xi-H(F(\xi))$ 在 M 上成立. □

逆映射定理常常以一种略微不同的陈述给出, 我们将其叙述为一个推论.

推论 11.1 在上面定理的假设下存在 β 的一个开邻域 U 使得 H 在 U 上为单射, $N=H(U)$ 在 V 中为开集, 而且 H^{-1} 在 N 上连续可微 (见图 3.11).

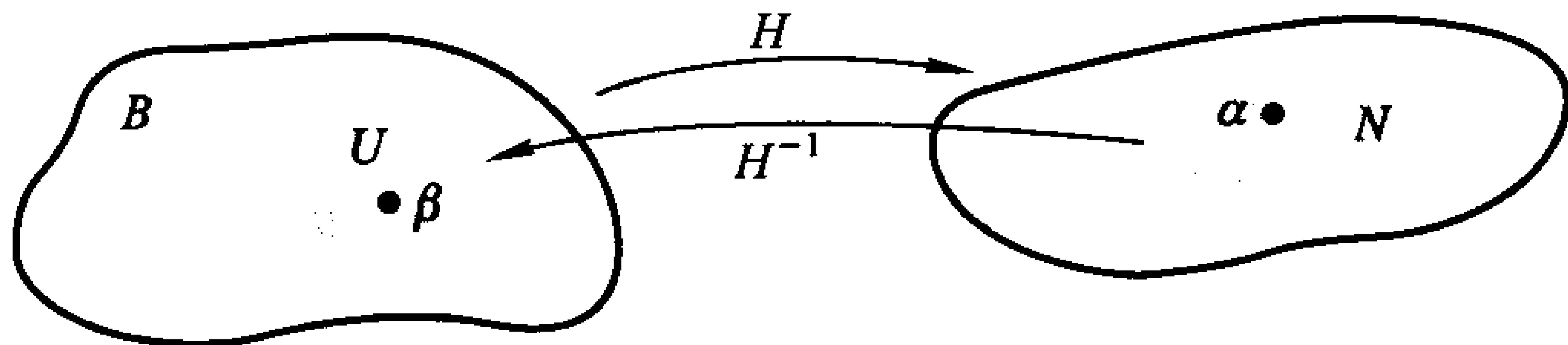


图 3.11

证明 留此证明为一个习题. □

在实用中我们常常不得不应用这些定理的笛卡儿式的陈述. 同学们应该肯定能把这些写出来, 但我们无论怎样也要叙述它们; 先从较简单的逆映射定理开始.

定理 11.4 假设我们有 n 个连续可微的 n 个实变量的实函数 $G_i(y_1,\cdots,y_2),i=1,\cdots,n$, 它们定义于 $\mathbb{R}^n$ 中点 $\boldsymbol{b}$ 的邻域 B 上, 又假设雅可比行列式

$$\frac{\partial(G_1,\cdots,G_n)}{\partial(y_1,\cdots,y_n)}(\boldsymbol{b})$$

非零. 于是存在 $\mathbb{R}^n$ 中 $\boldsymbol{a}=G(\boldsymbol{b})$ 的一个球 M 和一个惟一确定的 n 联组

$\boldsymbol{F} = \prec F_1, \cdots, F_n \succ$, 其中对所有 i F_i 为定义在 M 上的连续可微的实函数, 使得 $\boldsymbol{F}(\boldsymbol{a}) = \boldsymbol{b}$ 和 $G(F(\boldsymbol{x})) = \boldsymbol{x}$ 在 M 上成立. 那就是说, 对所有 $x \in M$ 及 $i = 1, \cdots, n, G_i(F_1(x_1, \cdots, x_n), \cdots, F_n(x_1, \cdots, x_n)) = x_i$.

例如, 如果 $\boldsymbol{x} = \prec y_1^3 + y_2^3, y_1^2 + y_2^2 \succ$, 则在点 $b = \prec 1, 2 \succ$ 我们有

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \det \begin{bmatrix} 3y_1^2, & 3y_2^2 \\ 2y_1, & 2y_2 \end{bmatrix} \Bigg|_{\prec 1,2 \succ}$$

$$= \det \begin{bmatrix} 3, & 12 \\ 2, & 4 \end{bmatrix} = -12 \neq 0,$$

因此不用去显式地求解我们便知道 $\boldsymbol{y}$ 有惟一用 $\boldsymbol{x}$ 表示的解, 它在 $\boldsymbol{x} = \prec 1^3 + 2^3, 1^2 + 2^2 \succ = \prec 9, 5 \succ$ 附近定义. 读者会发现真的不可能解出 $\boldsymbol{y}$, 这是因为他会很快发现必须要解一个 6 次的多项式方程. 这清楚地显示出此定理的威力: 我们保证了一个映射的存在性而要把它显式地解出来如果不是不可能的话也是非常困难的.(但是在下章中我们将发现一个迭代过程, 可以对逆映射近似到我们想要的那样靠近.)

我们在这里所说过的一切更加能应用于隐函数定理, 现在我们叙述它的笛卡儿形式.

定理 11.5 假设我们有定义于 $\mathbb{R}^{n+m}$ 中一个开子集 $A \times B$ 上的 $n+m$ 个实变量的 m 个连续可微的实函数 $G_i(\boldsymbol{x}, \boldsymbol{y}) = G_i(x_1, \cdots, x_n, y_1, \cdots, y_m)$ 以及一个 $(n+m)$ 联组 $\prec \boldsymbol{a}, \boldsymbol{b} \succ = \prec a_1, \cdots, a_n, b_1, \cdots, b_m \succ$ 使得对 $i = 1, \cdots, m$ 有 $G_i(\boldsymbol{a}, \boldsymbol{b}) = 0$, 并使得雅可比行列式

$$\frac{\partial(G_1, \cdots, G_m)}{\partial(y_1, \cdots, y_m)}(\boldsymbol{a}, \boldsymbol{b})$$

非零. 则存在 $\mathbb{R}^n$ 中关于 $\boldsymbol{a}$ 的一个球 M 和一个定义于 M 上的连续可微实函数 $F_j(x) = F_j(x_1, \cdots, x_n)$ 的一个惟一的 m 联组 $\boldsymbol{F} = \prec F_1, \cdots, F_m \succ$ 使得 $\boldsymbol{b} = \boldsymbol{F}(\boldsymbol{a})$ 及在 M 上 $G_i(\boldsymbol{x}, \boldsymbol{F}(\boldsymbol{x})) = 0, i = 1, \cdots, m$. 那就是说, $\boldsymbol{b}_i = F_i(a_1, \cdots, a_n), i = 1, \cdots, m$, 和 $G_i(x_1, \cdots, x_n; F_1(x_1, \cdots, x_n), \cdots, F_m(x_1, \cdots, x_n)) = 0$ 对所有 M 中 $\boldsymbol{x}$ 成立, 并且 $i = 1, \cdots, m$.

例如, 方程

$$x_1^3 + x_2^3 - y_1^2 - y_2^2 = 0,$$

$$x_1^2 - x_2^2 - y_1^3 - y_2^3 = 0$$

可以在 $\prec \boldsymbol{x}, \boldsymbol{y} \succ = \prec 1, 1, 1, -1 \succ$ 附近解出用 $\boldsymbol{x}$ 表示的惟一的 $\boldsymbol{y}$, 因为它们在那个点成立而且因为

$$\frac{\partial(G_1, G_2)}{\partial(y_1, y_2)} = \det \begin{bmatrix} -2y_1, & -2y_2 \\ -3y_1^2, & -3y_2^3 \end{bmatrix} = 6(y_1 y_2^2 - y_2 y_1^2)$$

在那里的值为 12. 当然, 我们只是表明解函数存在而不是我们能显式地给出它们.

习题

11.1 证明 $\prec x, y \succ \mapsto \prec e^x + e^y, e^x + e^{-y} \succ$ 在任意点 $\prec a, b \succ$ 都是局部可逆的, 并计算此逆映射的雅可比矩阵.

11.2 通过雅可比矩阵的计算证明 $\prec u, v \succ \mapsto \prec e^u + e^v, e^u - e^v \succ$ 在 $\mathbb{R}^2$ 中任意点局部可逆. 经简易计算这个局部逆知道在此时整个映射是可逆的. 做出这个计算, 并计算此逆映射的雅可比矩阵. 验证这两个矩阵在适当的点上是互逆的.

11.3 证明从 $\mathbb{R}^3$ 到 $\mathbb{R}^3$ 的映射 $\prec x, y, z \succ \mapsto \prec \sin x, \cos y, e^z \succ$ 关于 $\prec 0, \pi/2, 0 \succ$ 是局部可逆的. 证明

$$\prec x, y, z \succ \mapsto \prec \sin(x + y + z),\ \cos(x - y + z),\ e^{(x+y-z)} \succ$$

关于 $\prec \pi/4, -\pi/4, 0 \succ$ 是局部可逆的.

11.4 表达上面习题中的第二个映射为两个映射的复合, 并以第二种方式得到你的答案.

11.5 设 $F: \prec x, y \succ \mapsto \prec u, v \succ$ 为由 $u = x^2 + y^2, v = 2xy$ 定义的从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的映射. 计算 F 的逆 G, 并仔细地给出 G 的定义域和值域. 有多少逆映射? 计算 F 在 $\prec 1, 2 \succ$ 以及 G 在 $\prec 5, 4 \succ$ 的雅可比矩阵, 并把它们相乘以证明 F 和 G 互逆.

11.6 现考虑映射 $F: \prec x, y \succ \mapsto \prec x^3, y^3 \succ$. 证明 $dF_{\prec 0,0 \succ}$ 为奇异的但此映射有一个逆 G. 我们对 G 在原点的可微性得出什么结论?

11.7 定义 $F: \mathbb{R}^2 \to \mathbb{R}^2$ 为 $\prec x, y \succ \mapsto \prec e^x \cos y, e^x \sin y \succ$. 证明 F 关于每点都是局部可逆的.

11.8 定义 $F: \mathbb{R}^3 \to \mathbb{R}^3$ 为 $\boldsymbol{x} \mapsto \boldsymbol{y}$, 其中

$$y_1 = x_1 + x_2^2 + (x_3 - 1)^4, \quad y_2 = x_1^2 + x_2 + (x_3^3 - 3x_3), \quad y_3 = x_1^3 + x_2^2 + x_3.$$

证明 $\boldsymbol{x} \mapsto \boldsymbol{y} = F(\boldsymbol{x})$ 关于 $\boldsymbol{x} = \prec 0, 0, 1 \succ$ 为局部可逆的.

11.9 对一个函数 $f:\mathbb{R}\to\mathbb{R}$, 证明关于一点 a 在 df_a 非异时的局部可逆性远较一般情形简单. 先指出 f 在 a 的雅可比矩阵为数 $f'(a)$. 因此我们假定 $f'(x)$ 在 a 的一个邻域中连续并且 $f'(a)\neq 0$. 证明 f 在关于 a 的一个区间中是严格递增 (或递降) 的, 现在可以完成定理证明了.(见习题 1.12.)

11.10 证明方程

$$t^2+x^3+y^3+z^3=0,\quad t+x^2+y^2+z^2=4,\quad 1+x+y+z=0$$

在 $\prec t,x,y,z\succ=\prec 0,-1,1,0\succ$ 附近有可微解 $x(t),y(t),z(t)$.

11.11 证明方程

$$e^x+e^{2y}+e^{3u}+e^{4v}=4,\quad e^x+e^y+e^u+e^v=4$$

可以在点 $\prec 0,0,0,0\succ$ 附近惟一地解出用 x 和 y 表达的 u 和 v.

11.12 设 S 为方程

$$xz+\sin(xy)+\cos(xz)=1$$

在 $\mathbb{R}^3$ 中的图像. 确定在 $(0,1,1)$ 的邻域中 S 是否是下面中任一形式的可微函数的图像:

$$z=f(x,y),\quad x=g(y,z),\quad y=h(x,z).$$

11.13 已知从 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的函数 f 和 g 使得 $f(a,b,c)=0$ 和 $g(a,b,c)=0$, 写出关于 f 和 g 的偏导数的条件以保证一对惟一的可微函数 $y=h(x)$ 和 $z=k(x)$ 存在, 它们满足

$$h(a)=b,\quad k(a)=c,$$

和

$$f(x,y,z)=f(x,h(x),k(x))=0.$$

在 $\prec a,b,c\succ$ 附近 $g(x,y,z)=g(x,h(x),k(x))=0$.

11.14 设 $G(\xi,\eta,\zeta)$ 为从 $V=\prod_1^3 V_i$ 到 W 的一个连续可微映射使得 $dG^3_{\boldsymbol{\alpha}}:V_3\to W$ 可逆且 $G(\boldsymbol{\alpha})=G(\alpha_1,\alpha_2,\alpha_3)=0$. 证明存在在 $V_1\times V_2$ 中 $\prec\alpha_1,\alpha_2\succ$ 附近定义的一个惟一确定的函数 $\zeta=F(\xi,\eta)$, 使得 $G(\xi,\eta,F(\xi,\eta))=0$ 和 $F(\alpha_1,\alpha_2)=\alpha_3$. 又证明

$$dF^1_{\prec\xi,\eta\succ}=[-dG^3_{\prec\xi,\eta,\zeta\succ}]^{-1}[dG^1_{\prec\xi,\eta,\zeta\succ}],$$

其中 $\zeta=F(\xi,\eta)$.

11.15 设 $F(\xi,\eta)$ 是从 $V\times W$ 到 X 的一个连续可微函数, 并且假设 $dF^2_{\prec\alpha,\beta\succ}$ 可逆. 令 $\gamma=F(\alpha,\beta)$, 证明在 $X\times V\times W$ 中 $\prec\gamma,\alpha,\beta\succ$ 有一个乘积邻域 $L\times M\times N$ 和一个惟一的连续可微映射 $G:L\times M\to N$ 使得在 $L\times M$ 上, $F(\xi,G(\zeta,\xi))=\zeta$.

11.16 假定方程 $g(x,y,z)=0$ 能解出 z 为 x 和 y 的函数. 这意味着存在一个函数 $f(x,y)$ 使得 $g(x,y,f(x,y))=0$. 又假定一切均可微, 计算 $\partial z/\partial x$.

11.17 假设方程

$$g(x,y,z)=0 \quad 和 \quad h(x,y,z)=0$$

可解出 y 和 z 为 x 的函数. 计算 dy/dx.

11.18 假设 $g(x,y,u,v)=0$ 和 $h(x,y,u,v)=0$ 能被解出 u,v 为 x 和 y 的函数. 计算 $\partial u/\partial x$.

11.19 计算 dz/dx, 其中 $x^3+y^3+z^3=0$ 和 $x^2+y^2+z^2=1$.

11.20 如果 $t^3+x^3+y^3+z^3=0$ 和 $t^2+x^2+y^2+z^2=1$, 则 $\partial z/\partial x$ 含糊不清. 我们显然会想像这些变量中两个作为另外两个的函数, 可以是 z 为相关而 x 为独立. 但是是 t 还是 y 将是其他一个独立变量? 在每种假定下计算 $\partial z/\partial x$.

11.21 我们有 4 个 "物理变量" p,v,t 和 φ 使得它们中每一个是其余三个中任意二个的函数. 证明 $\partial t/\partial p$ 有二个完全不同的意义; 标出各个适当的函数并应用隐式求微分过程以把这两个意义之间的关系弄得清楚明晰.

11.22 再次来说明 "一维的" 情形在本质上是较为简单的理由. 设 G 是从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的一个连续可微映射使得 $G(a,b)=0$ 和

$$(\partial G/\partial y)(a,b)=G_2(a,b)>0.$$

说明存在正数 ε 和 δ 使得对 $(a-\delta,a+\delta)$ 中每个 c, 函数 $g(y)=G(c,y)$ 在 $[b-\varepsilon,b+\varepsilon]$ 上严格递增, 而且 $G(c,b-\varepsilon)<0<G(c,b+\varepsilon)$. 由介值定理 (习题 1.13) 得出结论是存在一个惟一的函数 $F:(a-\delta,a+\delta)\to(b-\varepsilon,b+\varepsilon)$ 使得

$$G(x,F(x))=0.$$

11.23 再次应用上面习题所使用的相同推理证明 F 为连续.

11.24 在反函数定理中证明了 $dF_\alpha=(dH_\beta)^{-1}$. 就是说 H 的逆的微分是 H 微分的逆. 用下列办法证明它:

(a) 应用隐函数定理;

(b) 由恒等式 $H(F(\xi))=\xi$ 直接计算.

11.25 还是在逆映射定理的行文中, 证明有一个在 A 中 β 的邻域 M 使得 $F(H(\eta))=\eta$ 在 M 上成立.(不要去做它而只是再次应用定理.)

11.26 继续与逆映射定理的上下文有关的讨论. 假定有 (下一章的) 结果即如果 dH_β^{-1} 存在, 则对充分靠近 β 的 ξ dH_β^{-1} 也存在. 证明存在 B 中 β 的一个开邻域 U 使得在 U 上 H 为单射, 而 $H(U)$ 是 V 中开集 N, 并且 H^{-1} 在 N 上连续可微.

11.27 运用习题 3.21 给出在逆映射定理行文中的关于利普希茨连续的局部逆存在性的直接证明.[提示: 用定理 7.4].

11.28 反函数可微性的直接证明比用隐函数定理证明更简单. 模仿定理 11.1 中有关内容的一般推理方式给出这样一个证明.

11.29 证明可按下面方式由反函数定理推导出隐函数定理: 令

$$H(\xi,\eta)=\prec \xi, G(\xi,\eta)\succ,$$

并证明 $dH_{\prec\alpha,\beta\succ}$ 具有方块图

$$\begin{array}{c|c} I & 0 \\ \hline dG^1 & dG^2 \end{array}.$$

然后从第一章的方块图结果证明 $dH^{-1}_{\prec\alpha,\beta\succ}$ 存在. 应用逆映射定理.

3.12 子流形和拉格朗日乘子

如果 V 和 W 为有限维空间, 分别具维数 n 和 m, 并且如果 F 是由 V 中开子集 A 到 W 的连续映射, 则 F(的图像) 是 $V\times W$ 的子集, 我们把它形象化为一种展开在 A 上的 "n 维曲面"S.(见第 10 节.) 我们称 F 为 $V\times W$ 中的一块 n 维贴片(patch). 更一般地, 如果 X 是个任意的 $(n+m)$ 维空向量空间, S 为 X 中一个子集. 如果有一个从 X 到一个乘积空间 $V\times W$ 的同构 φ, 使得 V 为 n 维而 $\varphi(s)$ 是 $V\times W$ 中一块贴片, 我们则称子集 S 为一块*n 维贴片*. 那就是说, 当 X 被想成是 $V\times W$ 时 S 便是在前面意义下的贴片. 这表明, 如果 π_1 是 $X=V\times W$ 到 V 上的投射, 则 $\pi_1(S)$ 是 V 中一个开子集 A, 而限制 $\pi_1\restriction S$ 为一对一的并有连续逆. 如果 π_2 是到 W 上的投射, 则 $F=\pi_2\circ(\pi_1\restriction S)^{-1}$ 是 A 到 W 的映射, 其在 $V\times W$ 中的图像为 S(当将 $V\times W$ 恒同于 X).

现在还有不是这类 "贴片" 曲面的重要曲面. 例如, 考虑 $\mathbb{R}^3$ 中单位球的表面:$S=\{\boldsymbol{x}:\sum_1^3 x_i^2=1\}$. S 显然是 $\mathbb{R}^3$ 中一个二维曲面, 不管我们如何努力将 $\mathbb{R}^3$ 表为一个直和它也不可能表示为一个图像. 但应该同样清楚, S 是相互重叠的曲面贴片的并集. 如果 α 为 S 上任一点, 则 α 在 $\mathbb{R}^3$ 任意的充分小邻域 N 交 S 为一块贴片; 我们取 V 为平行于在 α 的切平面的子空间, W 为通过 0 的垂线. 又, S 的这个性质是我们所说的子流形的完全恰当的定义.

$(n+m)$ 维向量空间 X 的一个子集 S 是 X 的一个*n 维子流形*指的是如果对 S 上每个 α 有一个在 X 中的邻域 N, 它与 S 的交为一个 n 维贴片.

我们说 S 是光滑的是指如果所有这些贴片 S_α 光滑, 那就是, 如果函数 $F: A \to W$ 在 $V \times W$ 中的图像为贴片 S_α(此时 X 被视为 $V \times W$), 而 F 对每个这样的贴片 S_α 为连续可微.

我们上面所考虑的球面是 $\mathbb{R}^3$ 中一个二维光滑子流形.

子流形经常被表现为映射的零集合. 例如, 上面的球面是由 $G(\boldsymbol{x}) = \sum_1^3 x_i^2 - 1$ 定义的从 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的映射 G 的零点集. 显然有一个保证这种零集合为子流形的条件很重要.

定理 12.1 设 G 为从 $(n+m)$ 维向量空间 X 中一个开子集 U 到 m 维向量空间 Y 的一个连续可微映射, 使得对 G 的零集合 S 上每个 α 有 dG_α 为满射. 于是 S 是 X 的一个 n 维子流形.

证明 选取 S 的任意点 γ. 因为 dG_γ 从 $(n+m)$ 维向量空间 X 到 m 维向量空间 Y 是满的, 我们知道 dG_γ 的零空间 V 有维数 n(第二章定理 2.4). 设 W 为 V 的任一补空间, 并想像 X 为 $V \times W$, 从而 G 成为两个向量变元的函数而 γ 是点 $\prec \alpha, \beta \succ$ 使 $G(\alpha, \beta) = 0$. $dG_{\prec\alpha,\beta\succ}$ 在 W 上的限制是 W 到 Y 的一个同构; 就是说, $(dG^2_{\prec\alpha,\beta\succ})^{-1}$ 存在. 因此由隐函数定理, 有一个 $\prec \alpha, \beta \succ$ 在 X 中的乘积邻域 $S_\delta(\alpha) \times S_\gamma(\beta)$, 它与 S 的交集是 $S_\delta(\alpha)$ 上一个函数的图像. 这就证明了定理. □

如果 S 是个光滑子流形, 则图像为 S 在 γ 附近的贴片的函数 F(这时将 X 适当地看作 $V \times W$) 是连续可微的, 因此 S 在 γ 有一个惟一确定 n 维切平面 M, 它在我们的 $\mathfrak{o}$ 近似的意义下在 γ 附近最紧密地贴合于 S. 如果 $\gamma = 0$, 此切平面是个 n 维子空间, 而在一般情形它是一个子空间 N 通过 γ 的平移. 我们称 N 为 S 在 γ 的切空间; 它的元素正好是 X 中切于通过 γ 所作参数弧的向量. 以后我们将要做的是独立于到一个向量空间的任何嵌入而描述一个 n 维流形 S. 对 S 在一点 γ 的切空间概念仍然是非常珍贵的, 但我们将不能把它形象化为承载 S 的一个空间 X 的实际切平面. 取而代之, 我们不得不在某种程度上构造出在 γ 切于 S 的向量空间.

定理 10.2 提供了思考的线索; 该定理告诉我们说, 如果 S 作为子流形嵌入到一个向量空间 X, 则每个在 γ 切于 S 的向量可以表示为 S 中某条光滑曲线在 γ 的惟一切向量. 从 S 中通过 γ 的光滑曲线的集合到在 γ 的切空间的这个映射不是单射的; 显然, 不同的曲线可以在 γ 互相相切, 故在那里有相同的切向量. 因此, 在 S 中对应于在 γ 的一个切向量的对象是一个通过 γ 的光滑曲线的等价

类, 而这事实上将是我们对一般流形的切向量的定义.

子流形的概念能让我们以一种简洁的方式考虑经典的"约束"极大值问题. 我们有有限维向量空间 X 的一个开子集 U, 一个定义于 U 上的可微实值函数 F, 和一个位于 U 中的子流形 S. 我们将假设子流形 S 是一个从 U 到一个向量空间 Y 的连续可微映射 G 的零集, 使得 dG_γ 对 S 上每个 γ 的满射. 我们想考虑当 γ 被"约束"在 S 中时极大化 (极小化)$F(\gamma)$ 的问题. 我们不能期望从令 $dF_\gamma = 0$ 通过求解 γ 来求出这种极大值点 γ. 这是因为 γ_0 可能不是 F 的判别点. 例如, 考虑 $g(\boldsymbol{x}) = \sum_1^3 x_i^2 - 1$ 为 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的函数, $F(\boldsymbol{x}) = x_2$. 这里由 $g = 0$ 定义的"曲面"是单位球面 $\sum_1^3 x_i^2 = 1$, 而在此球面上 F 在 $\prec 0,1,0 \succ$ 有极大值 1. 但是 F 是线性的, 故 $dF_\gamma = F$ 永不会是零变换. 以拉格朗日乘子法而知名的方案表明我们仍然可以通过对某个适当的函数 L 解 $dL_\gamma = 0$ 的方式求出这些被约束的判别点.

定理 12.2 *假设 F 在 S 上的点 γ 有极大值. 于是存在 Y^* 中一个泛函 l 使得 γ 为函数 $F - (l \circ G)$ 的判别点.*

证明 由隐函数定理, 我们可将 X 表示为 $V \times W$ 使得 S 在 γ 附近的邻域是由 V 中开集 A 到 W 的一个映射 H 的图像. 因此当把 F 和 G 表示为 $V \times W$ 上函数时, 我们有: 在 $\gamma = \prec \alpha, \beta \succ$ 附近 $G(\xi, \eta) = 0$ 当且仅当 $\eta = H(\xi)$, 并且 $F(\xi, \eta)$ 在此零曲面上的限制因而是由 $K(\xi) = F(\xi, H(\xi))$ 定义的函数 $K : A \to \mathbb{R}$. 按假定, α 是这个函数的一个判别点. 因此

$$0 = dK_\alpha = dF^1_{\prec\alpha,\beta\succ} + dF^2_{\prec\alpha,\beta\succ} \circ dH_\alpha.$$

从等式 $G(\xi, H(\xi)) = 0$ 我们也得到

$$0 = dG^1_{\prec\alpha,\beta\succ} + dG^2_{\prec\alpha,\beta\succ} \circ dH_\alpha.$$

由于 $dG^2_{\prec\alpha,\beta\succ}$ 是可逆的, 我们可以从第二个方程中解出 dH_α 并将它代入第一个方程, 然后为简明起见去掉下标我们得到

$$dF^1 - dF^2 \circ (dG^2)^{-1} \circ dG^1 = 0.$$

设 $l \in Y^*$ 为泛函 $dF^2 \circ (dG^2)^{-1}$. 于是我们有 $dF^1 = l \circ dG^1$ 并由定义, $dF^2 = l \circ dG^2$. 对第一个方程 (从右边) 复合 $\pi_1 : V \times W \to V$, 第二个复合 π_2 并相加, 得到 $dF_{\prec\alpha,\beta\succ} = l \circ dG_{\prec\alpha,\beta\succ}$. 即 $d(F - l \circ G) = 0$. □

至此我们一点也没有解释过“拉格朗日乘子”这个词. 这来自此定理的笛卡儿表达形式, 其中我们有笛卡儿空间 $\mathbb{R}^n$ 中开子集 U, $Y=\mathbb{R}^m$, $G=\prec g^1,\cdots,g^m\succ$ 而 l 在 Y^* 中有形式 $l_c: l(y)=\sum_1^m c_i y_i$. 于是 $F-l\circ G=F-\sum_1^m c_i g^i$ 以及 $d(F-l\circ G)_a=0$ 成为

$$\frac{\partial F}{\partial x_j}-\sum_1^m c_i\frac{\partial g^i}{\partial x_j}=0,\quad j=1,\cdots,n.$$

这 n 个方程连同 m 个方程 $G=\prec g^1,\cdots,g^m\succ=0$ 给出 $m+n$ 个未知量 $x_1,\cdots,x_n,c_1,\cdots,c_m$ 的 $m+n$ 个方程.

我们原先的那个平凡的例子将表明这是如何做出来的. 我们要极大化从 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的 $F(\boldsymbol{x})=x_2$, 它从属于约束 $\sum_1^3 x_i^2=1$. 这里的 $g(\boldsymbol{x})=\sum_1^3 x_i^2-1$ 也是由 $\mathbb{R}^3$ 到 $\mathbb{R}$ 的, 而我们的方法告诉我们去寻找一个从属于 $g=0$ 的 $F-cg$ 的判别点. 我们的方程组是

$$0-2cx_1=0,$$

$$1-2cx_2=0,$$

$$0-2cx_3=0,$$

$$\sum_1^3 x_i^2=1.$$

第一个说 $c=0$ 或 $x_1=0$，而第二个意味着 c 不能为 0. 因此 $x_1=x_3=0$, 而第 4 个方程表明了 $x_2=\pm 1$.

另一个我们问题的例子是极小化一个长方体表面面积 $A=\alpha(xy+yz+zx)$, 它从属于固定体积的约束, $xyz=V$. 定理说极小点是 $A-\lambda V$ 的一个判别点, 其中 λ 为某个常数, 从而令此函数的微分为 0, 我们得到方程

$$2(y+z)-\lambda yz=0,$$

$$2(x+z)-\lambda xz=0,$$

$$2(x+y)-\lambda xy=0,$$

连同约束条件

$$xyz=V.$$

前三个方程表示 $x=y=z$. 最后一个给出了公共值 $V^{1/3}$.

*3.13 函数相关性

大体上说, 这个问题是这样的: 如果我们有一组连续函数, 它们都定义于某个开集 A 上, 我们如何能够说出是否它们中某一些是其他那些的函数? 举例说, 如果我们有三个实连续函数 f_1, f_2 和 f_3, 我们如何能说出它们中一个是否是另外两个的函数, 譬如 f_3 是 f_1 和 f_2 的函数, 它意味着存在一个二元函数 $g(x,y)$ 使得 $f_3(t) = g(f_1(t), f_2(t))$ 对在公共定义域 A 中所有的 t 所立? 如果发生这种情形, 我们说 f_3*函数相关于*f_2 和 f_3. 它与问下面的问题几乎是完全相同的: 什么时候映射 $F: t \mapsto \prec f_1(t), f_2(t), f_3(t) \succ$ 的值域 S 是 $\mathbb{R}^3$ 中一个二维子流形? 然而这些问题中有值得注意的差别. 如果 f_3 函数相关于 f_1 和 f_2, 则 F 的值域确实位于 $\mathbb{R}^3$ 的一个二维子流形上面, 即 g 的图像上. 但是不能保证它自身形成一个二维子流形. 例如, f_2 和 f_3 都会同时函数相关于 f_1, $f_2 = g \circ f_1$, 和 $f_3 = h \circ f_1$, 这时 F 的值域位于 $\mathbb{R}^3$ 中的曲线 $\prec s, g(s), h(s) \succ$ 上, 它是个一维子流形. 从相反的方向说, 可以 F 的值域是二维子流形 M 而 f_3 并非函数依赖于 f_2 和 f_1. 这种情形下我们能说的全部是,*局部地*函数 $\{f_i\}_1^3$ 中的一个是另外两个的函数, 这是因为*局部地*M 是个曲面贴片; 这里采用了上一节的语言, 但是如果在弯曲的曲面 M 上稍微移开一些到另一个点的邻域, 我们可能会不得不解出这些函数中的另一个. 不过如果 $M = F$ 的值域是一个二维流形的子集, 那么说函数 $\{f_i\}_1^3$ 是*函数相关*是合乎道理的, 这让我们去考察这个更自然的概念.

如果我们假定 $F = \prec f_1, f_2, f_3 \succ$ 为连续可微, 并在 A 中某点 α, dF_α 的秩为 3, 于是隐函数定理意味着 $F[A]$ 包含了 $\mathbb{R}^3$ 中点 $F(\alpha)$ 的一个整个的球. 因此对 $M = F$ 的值域位于 $\mathbb{R}^3$ 中一个二维子流形的必要条件是 dF_α 的秩处处小于 3. 我们将看到, 事实上如果 dF_α 的秩对所有的 α 都为 2, 则 $M = F$ 的值域实质上是个二维流形.(我们在后面将要解释, 这仍然有一点小小的难点.) 我们的工具是隐函数定理及下面的这个定理, 即 T 的秩是 T 的 "下半连续" 函数; 这个定理完全可以出现得更早些.

定理 13.1 *设 V 和 W 为有限维向量空间, 它按某种方式被赋范. 于是对* $\mathrm{Hom}(V, W)$ *中任意 T, 存在一个 ε 使得*

$$\|S - T\| < \varepsilon \ \Rightarrow \ S\text{的秩} \geqslant T\text{的秩}.$$

证明 设 T 有零空间 N 和值域 R, 并设 X 为 N 在 V 中的任意一个补空间. 于是 T 在 X 上的限制是到 R 的一个同构, 从而以某个正数 m 下有界.(由定理 4.2, 它的从 R 到 X 的逆以某个 b 有界, 我们令 $m = 1/b$.) 于是如果

$\|S-T\|<m/2$, 便得到 S 在 X 以 $m/2$ 下有界, 这是由于不等式

$$\|T(\alpha)\|\geqslant m\|\alpha\| \quad 和 \quad \|(S-T)(\alpha)\|\leqslant (m/2)\|\alpha\|$$

一起蕴涵了 $\|S(\alpha)\|\geqslant (m/2)\|\alpha\|$. 特别, S 是 X 上的单射, 故 S 的秩 $=d(S$ 的值域$)\geqslant d(x)=d(R)=T$ 的秩. □

我们现在能够证明一般性的局部定理了.

定理 13.2 *设 V 和 W 为有限维空间, r 是小于 W 的维数的一个整数, 并设 F 从开子集 $A\subset V$ 到 W 的一个连续微分映射使得 dF_γ 的秩 $=r$ 对所有 A 中的 γ 成立. 于是 A 中每个 γ 有一个邻域 U 使得 $F(U)$ 是 W 的一个 r 维贴片子流形.*

证明 对 A 中一个固定的 γ, 设 V_1 和 Y 分别为 dF_γ 的零空间和值域, 设 V_2 为 V_1 在 V 中的一个补空间并将 V 看作 $V_1\times V_2$. 于是 F 成为一个二元函数 $F(\xi,\eta)$, 而且如果 $\gamma=\prec\alpha,\beta\succ$, 则 $dF^2_{\prec\alpha,\beta\succ}$ 是 V_2 到 Y 的一个同构. 此刻我们已经可以选取分解 $W=W_1\bigoplus W_2$, 关于它 $F[A]$ 将会是一个图像 (局部地). 我们简单地选取任何一个直和分解 $W=W_1\bigoplus W_2$ 使得 W_2 是 $Y=(dF_{\prec\alpha,\beta\succ}$ 的值域) 的补空间. 因此 W_1 可以是 Y 但也不必如此. 设 P 为 W 沿 W_2 到 W_1 的射影. 由于 Y 是 P 的零空间的一个补空间, 我们便知道 $P\restriction Y$ 是 Y 到 W_1 的同构. 特别, W_1 是 r 维的, 并且

$$P\circ dF_{\prec\alpha,\beta\succ}的秩=r.$$

另外, 这也是关键之处, P 对充分靠近 $\prec\alpha,\beta\succ$ 的所有 $\prec\xi,\eta\succ$ 是一个从 $dF_{\prec\xi,\eta\succ}$ 的值域到 W_1 的同构. 因为上面的秩定理意味着在 $\prec\alpha,\beta\succ$ 的某个邻域有 $P\circ dF_{\prec\xi,\eta\succ}$ 的秩 $\geqslant P\circ dF_{\prec\alpha,\beta\succ}$ 的秩 $=r$. 另一方面, $P\circ dF_{\prec\xi,\eta\succ}$ 的值域包含在 P 的值域即 W_1 中, 故 $P\circ dF_{\prec\xi,\eta\succ}\leqslant r$. 因此 $P\circ dF_{\prec\xi,\eta\succ}$ 的秩 $=r$ 对 $\prec\alpha,\beta\succ$ 附近的 $\prec\xi,\eta\succ$ 成立, 并且由假设条件 $dF_{\prec\xi,\eta\succ}$ 的秩 $=r$, 我们看出 P 在任何这样的 $dF_{\prec\xi,\eta\succ}$ 的值域上都是一个同构.

现定义 $H:W_1\times A\ \to\ W_1$ 为映射

$$\prec\zeta,\xi,\eta\succ\ \mapsto\ P\circ F(\xi,\eta)-\zeta.$$

如果 $\mu=P\circ F(\alpha,\beta)$, 则 $dH^3_{\prec\mu,\alpha,\beta\succ}=P\circ dF^2_{\prec\alpha,\beta\succ}$, 它是从 V_2 到 W_1 的一个同构. 因此由隐函数定理知道存在 $\prec\mu,\alpha,\beta\succ$ 的邻域 $L\times M\times N$ 和一个惟一

确定的连续可微的从 $L \times M$ 到 N 的映射 G, 使得在 $L \times M$ 上

$$H(\zeta, \xi, G(\zeta, \xi)) = 0,$$

即在 $L \times M$ 上

$$\zeta = P \circ F(\xi, G(\zeta, \xi)).$$

我们剩下来的论证在于证明 $F(\xi, G(\zeta, \xi))$ 只是 ζ 的函数. 我们先对上面方程作关于 ξ 的微分, 得到

$$0 = P \circ (dF^1 + dF^2 \circ G^2) = P \circ dF \circ \prec I, dG^2 \succ .$$

像上面所注意到的, P 在 $dF_{\prec \xi, \eta \succ}$ 的值域上对所充分靠近 $\prec \alpha, \beta \succ$ 的 $\prec \xi, \eta \succ$ 是一个同构, 并且如果我们假定 $L \times M$ 也被取成充分小使这个性质成立, 则上面的方程意味着对所有 $\prec \xi, \eta \succ \in L \times M$ 有

$$dF_{\prec \xi, \eta \succ} \circ \prec I, dG^2 \succ = 0.$$

但这正是说 $F(\xi, G(\zeta, \xi))$ 关于 ξ 的偏微分恒为 0, 从而

$$F(\xi, G(\zeta, \xi))$$

是一个只含 ζ 的连续可微函数 K:

$$F(\xi, G(\zeta, \xi)) = K(\zeta).$$

由于 $\eta = G(\zeta, \xi), \zeta = P \circ F(\xi, \eta)$, 从而我们有 $F(\xi, \eta) = K(P \circ F(\xi, \eta))$ 或者

$$F = K \circ P \circ F,$$

而且它在 $M \times N$ 中满足 $P \circ F(\xi, \eta) \in L$ 的这些点 $\prec \xi, \eta \succ$ 组成的开集 U 上成立. 如果我们把 W 想成是 $W_1 \times W_2$, 那么 F 和 K 是有序函数偶对:$F = \prec F^1, F^2 \succ$ 和 $K = \prec l, k \succ$, P 是映射 $\prec \zeta, \nu \succ \mapsto \zeta$, 而上面方程的第二个分量便是

$$F^2 = k \circ F^1.$$

因为 $F^1[U] = P \circ F[U] = L$, 上面方程说 $F[U]$ 是从 L 到 W_2 的映射 k 的图像. 此外, 因 L 是 r 维向量空间 W_1 的开子集, 从而 $F[U]$ 是 $W = W_1 \times W_2$ 中一个 r 维贴片流形. □

上面的定理包括了我们原来对函数相关性所提问题的回答.

推论 13.1 设 $F=\{f^i\}_1^m$ 为定义于赋范空间 V 中开子集 A 上的连续可微实函数的 m 联组, 又假定 dF_α 在 A 上的秩为常数 r, r 小于 m. 于是 A 中任意一点 γ 有一邻域 U, 在它上面函数中 $m-r$ 个函数相关于剩下的 r 个函数.

证明 由假设, $dF_\gamma=\prec df_\gamma^1,\cdots,df_\gamma^m\succ$ 的值域 Y 是 $\mathbb{R}^m$ 中一个 r 维子空间. 因此我们能对一个补子空间 W_2 选取标准基元素 $\{\delta^i\}$ 使其中 $m-r$ 个为其基, 而且我们不妨对函数 f^i 重新编号使得它们为 $\delta^{r+1},\cdots,\delta^m$. 于是 $\mathbb{R}^m$ 到 $\mathbb{R}^r=L(\delta^1,\cdots,\delta^r)$ 上的射影 P 是一个从 Y 到 $\mathbb{R}^r$ 的一个同构 (由于 Y 是其零空间的补), 并且由上面定理知道存在 γ 的一个邻域 U, 在它上面 $(I-P)\circ F$ 是 $P\circ F$ 的一个函数 k. 但这正好是说 $\prec f^{r+1},\cdots,f^m\succ=k\circ\prec f^1,\cdots,f^r\succ$, 即 k 是一个取值 $(m-r)$ 联组的函数, $k=\prec k^{r+1},\cdots,k^m\succ$, 并对 $j=r+1,\cdots,m$, $f^j=k^j\circ\prec f^1,\cdots,f^r\succ$. □

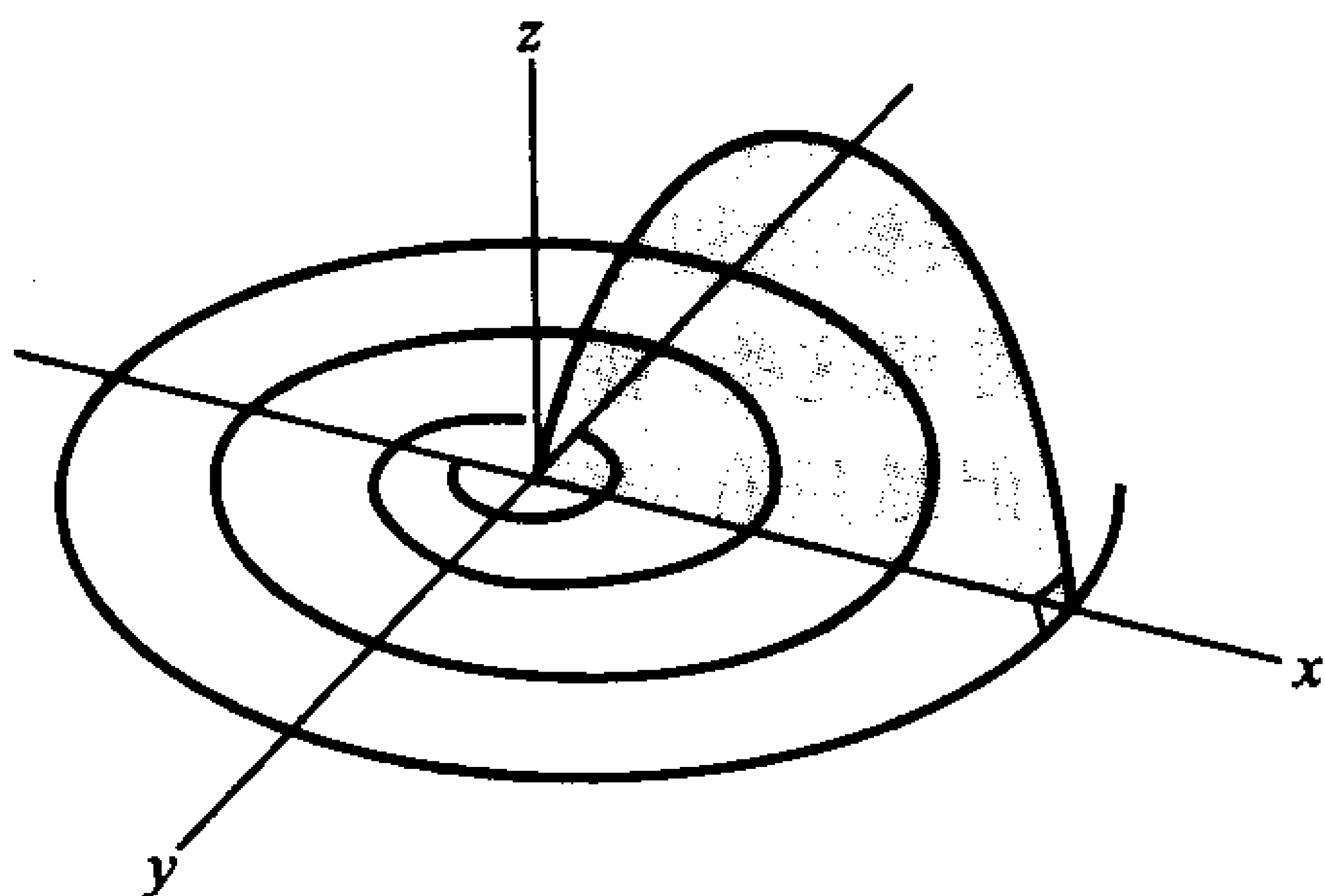


图 3.12

在本节前面部分我们曾提到, 在下结论说如果 F 是从 V 的开子集 A 到 W 的一个连续可微映射, 并且其微分具有小于 $d(W)$ 的常数秩 r, 则 $S=F$ 的值域是 W 的一个 r 维子流形时这里是有些难点的. 其不足之处可叙述于后. X 的子流形 S 的定义需要在 S 的每个点上有一个在 X 中的邻域, 使它与 S 的交集是个贴片. 在我们前面的情形中, 我们所能断定的是如果 β 是 S 中的点, 则在 A 中 $\beta=F(\alpha)$ 对某个 α 成立, 而且 α 有一个邻域 U 其在 F 下的像是个贴片. 但是这个像可以不是 β 在 S 中的整个邻域, 原因是 S 可以向自己弯曲回来以致于它侵入到 β 的每一个邻域中. 例如考虑按图 3.12 所提出的那个嵌入在 $\mathbb{R}^3$ 中的一维 Γ. 此曲线在 xz 平面开始沿着 z 轴进行, 弯曲过来, 当它到达 xy 平面时, 它开始

螺旋地趋向 xy 平面中的原点 (从 xz 平面进入 xy 平面处的改变点是个奇点, 但我们可以将其光滑化). 原点并不是具有与 Γ 相交出一个一维贴片的 $\mathbb{R}^3$ 中的邻域的那种点, 然而整条曲线是 $(-1,1)$ 在连续可微单射下的像.

我们没有任何困难地将 Γ 考虑为一个一维流形, 但是联系到它在 $\mathbb{R}^3$ 中的嵌入时, 一些东西就会出错了, 这时它已不是 $\mathbb{R}^3$ 中一个一维子流形.

*3.14 一致连续性和取函数为值的映射

在下一章我们将看到, 一个在有限维向量空间 V 的有界闭子集上定义的连续函数 F 必定是一致连续的. 这意味着对给出的 ε 存在一个 δ 使得

$$\|\xi-\eta\|<\delta \ \Rightarrow \ \|F(\xi)-F(\eta)\|<\varepsilon$$

对 F 定义域中所有的向量 ξ 和 η 成立.

关键之处在于 δ 只依赖于 ε 而不是像普通的连续性那样, 依赖于被断言具连续性的那个 "锚定" 的点. 这是个非常重要的性质. 在本节我们将看到它支撑着一类定理, 这类定理中一个点映射被升级为一个取函数为值的映射, 同时点映射的性质蕴涵了取函数为值的映射的相应性质. 这些定理有很强的应用性, 我们将在第 15 节和第六章第 1 节看到这些. 我们在这里立刻要给出的一个应用是关于在积分符号下的微分定理. 但是只有定理 14.3 会在本书后面被用到.

首先假设 $F(\xi,\eta)$ 是从乘积开集 $M\times N$ 到赋范线性空间 X 的一个有界连续函数. 使 η 固定, 我们有函数 $f_\eta(\xi)=F(\xi,\eta)$, 它是在 M 上的有界连续函数, 就是说, 是所有从 M 到 X 的有界连续函数的赋范线性空间 $Y=\mathfrak{Bc}(M,X)$ 中的一个元素. 这个函数也被表示为 $F(\cdot,\eta)$, 故 $f_\eta=F(\cdot,\eta)$. 我们假定在 Y 上用一致范数:

$$\|f_\eta\|=\mathrm{lub}\{\|f_\eta(\xi)\|:\xi\in M\}=\mathrm{lub}\{\|F(\xi,\eta)\|:\xi\in M\}.$$

定理 14.1 *在上面的行文中, 如果 F 为一致连续, 则映射 $\eta\mapsto f_\eta$(或 $\eta\mapsto F(\cdot,\eta)$)$:N\to Y$ 为连续, 事实上为一致连续.*

证明 对给出的 ε, 选取 δ 使得

$$\|\prec\xi,\eta\succ-\prec\mu,\nu\succ\|<\delta \ \Rightarrow \ \|F(\xi,\eta)-F(\mu,\nu)\|<\varepsilon.$$

取 $\mu=\xi$ 并重写上式右端, 我们有

$$\|\eta-\nu\|<\delta \ \Rightarrow \ \|f_\eta(\xi)-f_\nu(\xi)\|<\varepsilon$$

对所有 ξ 成立. 因此

$$\|\eta - \nu\| < \delta \ \Rightarrow\ \|f_\eta - f_\nu\|_\infty \leqslant \varepsilon.$$

我们已经证明如果一个二元函数为一致连续则用一般对偶原理从它得到的映射是连续的. 许多熟知的事实后面都有这种现象. 例如:

推论 14.1 如果 $F(x,y)$ 是在 $\mathbb{R}^2$ 中单位正方形 $[0,1]\times[0,1]$ 上的一个一致连续的实函数, 则 $\int_0^1 F(x,y)dx$ 是 y 的连续函数.

证明 映射 $y \mapsto \int_0^1 F(x,y)dx$ 是从 $\mathfrak{C}([0,1])$ 到 $\mathbb{R}$ 的有界线性映射 $f \mapsto \int_0^1 f$ 与从 $[0,1]$ 到 $\mathfrak{C}([0,1])$ 的连续映射 $y \mapsto F(\cdot,y)$ 的复合, 从而作为连续映射的复合仍是连续的. □

我们下面考虑前面的对偶性诱导映射的可微性.

定理 14.2 如果 F 是从赋范线性空间 $V\times W$ 的开的乘积集合 $M\times N$ 到赋范线性空间 X 的一个有界连续映射, 又如果 $dF^2_{\prec\alpha,\beta\succ}$ 存在, 并且是 $M\times N$ 上 $\prec\alpha,\beta\succ$ 的有界一致连续函数, 则 $\varphi:\eta\mapsto F(\cdot,\eta)$ 是从 N 到 $Y=\mathfrak{BC}(M,N)$ 的一个可微映射, 并且 $[d\varphi_\beta(\eta)](\xi) = dF^2_{\prec\xi,\beta\succ}(\eta)$.

证明 对给出的 ε, 我们由 dF^2 的一致连续性选取 δ, 使得对所有 $\xi\in M$

$$\|\mu - \nu\| < \delta \ \Rightarrow\ \|dF^2_{\prec\xi,\mu\succ} - dF^2_{\prec\xi,\nu\succ}\| < \varepsilon.$$

定理 7.4 的推论则表明

$$\|\eta\| < \delta \ \Rightarrow\ \|\Delta F^2_{\prec\xi,\beta\succ}(\eta) - dF^2_{\prec\xi,\beta\succ}(\eta)\| \leqslant \varepsilon\|\eta\|$$

对所有的 $\xi\in M$, 所有 $\beta\in N$ 以及所有使从 β 到 $\beta+\eta$ 的线段均在 N 中的 η 成立. 固定 β, 并将上面不等式的右端重写. 这里是此证明的中心部分. 首先

$$\begin{aligned}\Delta F^2_{\prec\xi,\beta\succ}(\eta) &= F(\xi,\beta+\eta) - F(\xi,\beta)\\ &= [f_{\beta+\eta} - f_\beta](\xi) = [\varphi(\beta+\eta) - \varphi(\beta)](\xi)\\ &= [\Delta\varphi_\beta(\eta)](\xi).\end{aligned}$$

其次我们可验证如果 $\|dF^2_{\prec\mu,\nu\succ}\| \leqslant b$, 对于 $\prec\mu,\nu\succ\in M\times N$, 则由公式 $[T(\eta)](\xi) = dF^2_{\prec\xi,\beta\succ}(\eta)$ 定义的映射 T 是 $\mathrm{Hom}(W,Y)$ 中一个元素, 其范数最多为

b. 对此详细的验证当作一个习题留给读者. 最后展示的那个不等式现在具有的形式是

$$\|\eta\| < \delta \Rightarrow \|[\Delta\varphi_\beta(\eta) - T(\eta)](\xi)\| \leqslant \varepsilon\|\eta\|,$$

从而

$$\|\eta\| < \delta \Rightarrow \|\Delta\varphi_\beta(\eta) - T(\eta)\|_\infty \leqslant \varepsilon\|\eta\|.$$

这准确地说出映射 φ 在 β 可微且 $d\varphi_\beta = T$. □

事实上映射 φ 是连续可微的, 这可由按上面方式的稍许进一步的论证看出. 这个情况极其近似于定理 14.1 的一个应用.

关于积分符号下可微性的经典定理是上面定理的推论. 我们给出一个简单的情形. 注意, 如果 η 是个实变量 y, 则上面对 $d\varphi$ 的公式可以用弧的导数表示重写:

$$[\varphi'(b)](\xi) = \frac{\partial F}{\partial y}(\xi, b).$$

推论 14.2 *如果 $F(x,y)$ 为单位正方形 $[0,1]\times[0,1]$ 上的连续实函数, 并且如果 $\partial F/\partial y$ 存在且在此正方形上是一个一致连续函数, 则 $\int_0^1 F(x,y)dx$ 是 y 的可微函数, 其导数为 $\int_0^1 (\partial F/\partial y)(x,y)dx$.*

证明 映射 $T : y \mapsto \int_0^1 F(x,y)dx$ 是一个复合函数: 从 $\mathfrak{C}([0,1])$ 到 $\mathbb{R}$ 的有界线性映射 $f \mapsto \int_0^1 f(x)dx$ 复合上从 $[0,1]$ 到 $\mathfrak{C}([0,1])$ 的可微函数 $\varphi : y \mapsto F(\cdot, y)$, 因此由复合函数规则它是可微的. 于是定理 7.2 和有界线性映射的微分是自身这一事实给出了

$$T'(y) = \int_0^1 [\varphi'(y)](x)dx = \int_0^1 \frac{\partial F}{\partial y}(x,y)dx. \quad □$$

现在我们到了对我们来说最重要的情形, 即一个点到点的映射由复合生成一个函数到函数的映射的情形. 设 A 为赋范线性空间 V 中开集, S 为一任意集合, 并设 $\mathfrak{a}$ 为从 S 到 A 的有界映射的集合. 于是 $\mathfrak{a}$ 是赋范线性空间 $\mathfrak{B}(S,V)$ 的子集, 这里的 $\mathfrak{B}(S,V)$ 是 S 到 V 在一致范数下的所有有界函数集. 函数 $f \in \mathfrak{a}$ 为 $\mathfrak{a}$ 的内点当且仅当 f 到 $\mathfrak{a}$ 的边界的距离是个函数 δ, 因为这显然等价于说 $\mathfrak{a}$ 包含了在 $\mathfrak{B}(S,V)$ 的点 f 的一个球, 现设 g 为由 A 到赋范线性空间 W 的任一有界映射, 并设 $G : \mathfrak{a} \to \mathfrak{B}(S,W)$ 为复合上 g. 就是说, $h = G(f)$ 当且仅当 $f \in \mathfrak{a}$ 且 $h = g \circ f$. 我们可以同时考虑 G 的连续性和可微性, 但我们只做了可微性定理.

定理 14.3 设函数 $g:A\to W$ 在每点 $\alpha\in A$ 可微，又设 dg_α 是 α 的一个有界一致连续函数. 则由 $G(f)=g\circ f$ 定义的映射 $G:\mathfrak{a}\to\mathfrak{B}(S,W)$ 在 $\mathfrak{a}$ 的每个内点 f 可微，并且 $dG_f:\mathfrak{B}(S,V)\to\mathfrak{B}(S,W)$ 对所有 $s\in S$ 由

$$[dG_f(h)](s)=dg_{f(s)}(h(s))$$

定义.

证明 对给出的 ε, 选取 δ, 由 dg 的一致连续性, 使得

$$\|\alpha-\beta\|<\delta\ \Rightarrow\ \|dg_\alpha-dg_\beta\|<\varepsilon,$$

然后再次应用定理 7.4 的推论, 并假定从 α 到 $\alpha+\xi$ 的线段在 A 中, 则结论是

$$\|\xi\|<\delta\ \Rightarrow\ \|\Delta g_\alpha(\xi)-dg_\alpha(\xi)\|\leqslant\varepsilon\|\xi\|.$$

现在选择 $\mathfrak{a}$ 中一个任意的固定的内点 f, 并选取 $\delta'\leqslant\delta$ 使得 $B_{\delta'}(f)\subset\mathfrak{a}$. 于是对 $\mathfrak{B}(S,V)$ 中任意 h, 并对所有 $s\in S$ 有

$$\|h\|_\infty<\delta'\ \Rightarrow\ \|\Delta g_{f(s)}(h(s))-dg_{f(s)}(h(s))\|\leqslant\varepsilon\|h(s)\|.$$

定义映射 $T:\mathfrak{B}(S,V)\to\mathfrak{B}(S,W)$ 为 $[T(h)](s)=dg_{f(s)}(h(s))$. 于是上面所展示的不等式可重写为

$$\|h\|_\infty<\delta'\Rightarrow\|\Delta G_f(h)-T(h)\|_\infty\leqslant\varepsilon\|h\|_\infty.$$

这就是说, $\Delta G_f=T+\mathfrak{o}$. 因此在我们证明了 $T\in\mathrm{Hom}(\mathfrak{B}(S,V),\mathfrak{B}(S,W))$ 后就证好了.

首先, 我们有

$$\begin{aligned}(T(h_1+h_2))(s)&=dg_{f(s)}((h_1+h_2)(s))=dg_{f(s)}(h_1(s)+h_2(s))\\&=dg_{f(s)}(h_1(s))+dg_{f(s)}(h_2(s))\\&=(T(h_1))(s)+(T(h_2))(s).\end{aligned}$$

因此 $T(h_1+h_2)=T(h_1)+T(h_2)$, 齐性相似地得到. 其次, 如果 b 是 $\|dg_\alpha\|$ 在 A 上的一个界, 则 $\|T(h)\|_\infty=\mathrm{lub}\{\|T(h)(s)\|:s\in S\}\leqslant\mathrm{lub}\{\|dg_{f(s)}\|\cdot\|h(s)\|:s\in S\}\leqslant b\|h\|_\infty$. 因此 $\|T\|\leqslant b$. 证完. □

在上面的情形中, 如果 g 从 $A \times U$ 到 W, 故 $G(f)$ 是由 $h(t) = g(f(t), t)$ 给出的函数 h, 于是除了定理是关于 dg' 而非 dg 外没有任何改变. 另外, 如果 V 是乘积空间 $V_1 \times V_2$, 使得 f 具有形式 $\prec f_1, f_2 \succ$ 和 $[G(f)](t) = g(f_1(t), f_2(t), t)$, 于是关于偏微分的规则给了我们公式

$$[dG_f(h)](t) = dg^1_{f(t)}(h_1(t)) + dg^2_{f(t)}(h_2(t)).$$

*3.15 变分法

变分法的问题简单地说是某种类型的判别点问题, 但是在应用条件 $dF_\alpha = 0$ 的方式有一个独特的变化. 我们用证明它的标准定理中的一个来解释这个学科.

由于要解定义域为无穷维向量空间的约束极大值问题, 一个系统的讨论要从比拉格朗日乘子定理更为一般的形式出发. 但是就我们的目的而言, 注意到如果 S 是闭曲面 $M + \alpha$, 则 F 在 S 上的限制等价于在向量空间 M 上的一个新函数, 而它在 $\beta = \eta + \alpha \in S$ 的微分显然正好是 dF_β 在 M 上的限制这样的事实就足够了. 要求 β 对约束函数是个判别点因而只不过是要求 dF_β 在 M 上为零.

设 F 是定义于 $W \times W \times \mathbb{R}$(一个开子集) 上的一致连续的可微的三元实函数, 其中 W 是个赋范线性空间. 给出闭区间 $[a, b] \in \mathbb{R}$, 并设 V 是赋范线性空间 $\mathfrak{C}^1([a, b], W)$, 它由 $f : [a, b] \to W$ 并取 $\|f\|$ 为 $\|f\|_\infty + \|f'\|_\infty$ 的光滑弧组成. 问题是极大化 (非线性) 泛函 $G(f) = \int_a^b F(f(t), f'(t), t)dt$, 它从属于约束 $f(a) = \alpha$ 和 $f(b) = \beta$. 就是说, 我们只考虑 W 中具固定端点 α 和 β 的光滑弧, 而我们要找那条使此积分极大 (或极小) 的从 α 到 β 的弧. 现在我们可以证明 G 是从 V(的一个开子集) 到 $\mathbb{R}$ 的一个连续可微函数. 证明的最容易方法是设 X 为连续弧的空间 $\mathfrak{C}([a, b], W)$, 其范数为一致范数, 并首先考虑更一般的从 $X \times X$ 到 $\mathbb{R}$ 的泛函 K, 它由 $K(f, g) = \int_a^b F(f(t), g(t), t)dt$ 定义. 由定理 14.3 知, 被积函数映射 $\prec f, g \succ \mapsto F(f(\cdot), g(\cdot), \cdot)$ 从 $X \times X$ 到 $\mathfrak{C}([a, b])$ 是可微的, 它在 $\prec f, g \succ$ 的微分在 $\prec h, k \succ$ 的取值为函数

$$dF^1_{\prec f(t), g(t), t \succ}(h(t)) + dF^2_{\prec f(t), g(t), t \succ}(k(t)).$$

因为 $f \mapsto \int_a^b f(t)$ 是 $\mathfrak{C}$ 上的一个有界线性泛函, 它可微并等于它的微分. 因而复合函数规则意味着 K 可微且

$$dK_{\prec f, g \succ}(h, k) = \int_a^b [dF^1(h(t)) + dF^2(k(t))]dt,$$

其中在被积项的偏微分是取在点 $\prec f(t), g(t), t \succ$. 现在使 f' 存在并等于 g 的偶对 $\prec f, g \succ$ 形成 $X \times X$ 中一个闭子空间, 它同构于 V. 显然它们形成一个子空间, 但要看出它是闭的则需要第四章的对参数弧的积分理论, 因为它依赖于表达式 $f(t) = f(a) + \int_a^t f'(s)ds$ 和作为结果的范数不等式 $\|f(t) - f(a)\| \leqslant (t-a)\|f'\|_\infty$. 假定有这些结果, 我们则看出原来的泛函 G 恰好是 K 在此子空间 (同构于)V 的限制, 从而可微并有

$$dG_f(h) = \int_a^b [dF^1(h(t)) + dF^2(h'(t))]dt.$$

这个微分 dG_f 被称作 G关于 f 的第一变分.

弧 f 的固定端点 α 和 β 转而决定了 V 中一个闭的平面 P, 因为赋值映射 (坐标投影)$\pi_x : f \mapsto f(x)$ 为有界而 P 是超平面 $\pi_a = \alpha$ 和 $\pi_b = \beta$ 的交. 由于 P 是子空间 $M = \{f \in V : f(a) = f(b) = 0\}$ 的一个平移, 我们具约束的极大值方程为

$$dG_f(h) = \int_a^b [dF^1(h(t)) + dF^2(h'(t))]dt = 0$$

对 M 中所有 h 成立.

我们现在到了用变分法特殊技巧的时刻, 这个技巧被称作杜布瓦 - 雷蒙德 (Du Bois-Reymond) 引理.

为简明起见假设 $W = \mathbb{R}$. 则 F 是三个实变量的函数 $F(x, y, t)$, 偏微分等价于普通的偏导数, 而我们的判别点方程是

$$dG_f(h) = \int_a^b \left(\frac{\partial F}{\partial x} \cdot h + \frac{\partial F}{\partial y} \cdot h'\right) = 0.$$

如果我们用分部积分法到积分中的第一项, 并记住 $h(a) = h(b) = 0$, 则此方程成为

$$\int_a^b \left(\frac{\partial F}{\partial y} - \int \frac{\partial F}{\partial x}\right) g = 0,$$

其中 $g = h'$. 由于 h 是除去约束 $h(a) = h(b) = 0$ 外是任意的连续可微函数, 我们看到 g 是除去约束 $\int_a^b g(t)dt = 0$ 外的任意连续函数. 这就是说, $\frac{\partial F}{\partial y} - \int \frac{\partial F}{\partial x}$ 正交于线性泛函 $g \mapsto \int_a^b g(t)dt$ 的零空间 N. 由于一维空间 $N^\perp$ 显见是常值函数的集合,

$$\int_a^b g = 0 \Rightarrow \int_a^b Cg = 0,$$

我们的条件变成了

$$\frac{\partial F}{\partial y}(f(t),f'(t),t)=\int_0^t \frac{\partial F}{\partial x}(f(s),f'(s),s)ds+C.$$

这个方程特别表明其左端的函数可微. 这并不是很显见的, 因为我们只假定了 f' 为连续. 微分此方程, 我们最后得出的结论是, f 为映射 G 的一个判别点当且仅当它是微分方程

$$\frac{d}{dt}\frac{\partial F}{\partial y}(f(t),f'(t),t)=\frac{\partial F}{\partial x}(f(t),f'(t),t)$$

的解. 称此方程为变分问题的*欧拉方程*. 这是对未知函数 f 的一个常微分方程; 当所表示的导数计算出来后, 它具有形式

$$\frac{\partial^2 F}{\partial y^2}f''+\frac{\partial^2 F}{\partial y\partial x}f'+\frac{\partial^2 F}{\partial y\partial t}-\frac{\partial F}{\partial x}=0.$$

如果 W 不是 $\mathbb{R}$, 我们用分部积分法公式的一般形式 (用定理 6.3) 我们完全准确地得到同样的结果和一个比上述论证的更加精妙的形式. (见第四章的习题 10.14 和 10.15.) 那即是, 具固定端点 α 和 β 的光滑弧 f 是映射 $g\mapsto \int_a^b F(g(t),g'(t),t)dt$ 的判别点当且仅当它满足欧拉微分方程

$$\frac{d}{dt}dF^2_{\prec f(t),f'(t),t\succ}=dF^1_{\prec f(t),f'(t),t\succ}.$$

现在这是一个向量值的方程, 取值于 W^*. 如果 W 是有限维, 维数为 n, 则选择基可使 W^* 成为 $\mathbb{R}^n$, 而此向量方程等价于 n 个纯量方程

$$\frac{d}{dt}\frac{\partial F}{\partial y_i}(f(t),f'(t),t)=\frac{\partial F}{\partial x_i}(f(t),f'(t),t),$$

其中 F 现在是一个 $2n+1$ 实变量的函数

$$F(\boldsymbol{x},\boldsymbol{y},t)=F(x_1,\cdots,x_n,y_1,\cdots,y_n,t).$$

最后, 让我们看一看当 f 的端点不固定时对较简单的变分问题 ($W=\mathbb{R}$) 会发生什么. 现在判别点方程是 $dG_f(h)=0$ 对 V 中所有h 成立, 并且当我们分部积分时它成为

$$\frac{\partial F}{\partial y}\cdot h\Big]_a^b+\int_a^b\left(\frac{\partial F}{\partial x}-\frac{d}{dt}\frac{\partial F}{\partial y}\right)h=0$$

对 V 中所有的 h 成立. 本质上我们能按上面那样推理, 但更为精细一些, 得到的结果是, 一个函数 f 为判别点当且仅当它满足欧拉方程

$$\frac{d}{dt}\left(\frac{\partial F}{\partial y}\right) - \frac{\partial F}{\partial x} = 0$$

以及端点条件

$$\left.\frac{\partial F}{\partial y}\right|_{t=a} = \left.\frac{\partial F}{\partial y}\right|_{t=b} = 0.$$

这里只是对变分法的一个快速浏览, 感兴趣的读者可进一步查寻关于此学科的专著. 对我们考虑过的这种一般类型有着许许多多问题. 例如, 我们可以既不要求固定也不要求完全自由的端点但是要有隶属于约束条件的自由度. 我们在第十三章中的力学的变分方程的特殊情形将着手这个问题. 或许 f 再次为一个或更多变量的函数但积分会是多重的. 这时的欧拉方程成为一组未知函数 f 的偏微分方程. 最后, 对判别函数是这个积分的极大或极小值的充分条件仍是个问题. 它自然涉及到对泛函 G 的二阶微分的研究, 或者像在这门学科中所称的二阶变分的研究.

*3.16 二阶微分和判别点的分类

假设 V 和 W 是赋范线性空间, A 是 V 的开子集, 而 $F: A \to W$ 为连续可微映射. F 的一阶微分是从 A 到 $\mathrm{Hom}(V, W)$ 的连续映射 $dF: \gamma \mapsto dF_\gamma$. 现在我们要研究这个映射在点 α 的可微性. 按假定, 我们知道说 dF 在 α 可微的意思是什么. 由定义 $d(dF)_\alpha$ 是一个从 V 到 $\mathrm{Hom}(V, W)$ 的有界线性变换 T, 使得 $\Delta(dF)_\alpha(\eta) - T(\eta) = \mathfrak{o}(\eta)$. 就是说, $dF_{\alpha+\eta} - dF_\alpha - T(\eta)$ 是 $\mathrm{Hom}(V, W)$ 中对充分小的 η 其范数小于 $\varepsilon\|\eta\|$ 的一个元素. 我们令 $d^2F_\alpha = d(dF)_\alpha$, 重说一遍就是: $d^2F_\alpha = d^2F_\alpha(\cdot)$ 是从 V 到 $\mathrm{Hom}(V, W)$ 的一个线性映射. $d^2F_\alpha(\eta) = d^2F_\alpha(\eta)(\cdot)$ 是 $\mathrm{Hom}(V, W)$ 的一个元素, 而 $d^2F_\alpha(\eta)(\xi)$ 是 W 的一个向量. 又, 我们知道 d^2F_α 等价于一个有界双线性映射

$$\omega: V \times V \to W,$$

其定义为 $\omega(\eta, \xi) = d^2F_\alpha(\eta)(\xi)$.

向量 $d^2F_\alpha(\eta)$ 显见应该是 F 在 α 的某种二阶导数, 读者甚至会猜出它是沿方向 ξ 和 η 的混合导数.

定理 16.1 如果 $F: A \to W$ 为连续可微, 且二阶微分 d^2F_α 存在, 则对每个固定的 $\mu \in V$ 从 A 到 W 的函数 $D_\mu F: \gamma \mapsto D_\mu F(\gamma)$ 在 α 可微, 而且 $D_\nu(D_\mu F)(\alpha) = (d^2F_\alpha(\nu))(\mu)$.

证明 我们运用在 μ 赋值的映射 $\mathrm{ev}_\mu: \mathrm{Hom}(V, W) \to W$, 其定义为: 对 V 中固定的 μ, $\mathrm{ev}_\mu(T) = T(\mu)$. 它是一个有界线性映射. 于是

$$(D_\mu F)(\alpha) = dF_\alpha(\mu) = \mathrm{ev}_\mu(dF_\alpha) = (\mathrm{ev}_\mu \circ dF)(\alpha),$$

故函数 $D_\mu F$ 是复合 $\mathrm{ev}_\mu \circ dF$. 因为 ev_μ 为线性并且 $d(dF)_\alpha$ 存在, 故它在 α 可微. 因此 $(D_\nu(D_\mu F))(\alpha) = d(D_\mu F)_\alpha(\nu) = d(\mathrm{ev}_\mu \circ dF)_\alpha(\nu) = (\mathrm{ev}_\mu \circ d(dF)_\alpha)(\nu) = \mathrm{ev}_\mu[(d^2F_\alpha)(\nu)] = (d^2F_\alpha(\nu))(\mu)$. □

读者必定想起在上面的论证过程中, $D_\mu F$ 是函数 $(D_\mu F)(\cdot)$, 他或许宁愿使用这个符号, 计算如下.

$$\begin{aligned} D_\nu((D_\mu F)(\cdot))|_\alpha &= d((D_\mu F)(\cdot))_\alpha(\nu) = d(\mathrm{ev}_\mu \circ dF_{(\cdot)})_\alpha(\nu) \\ &= [\mathrm{ev}_\mu \circ d(dF_{(\cdot)})_\alpha](\nu) \\ &= \mathrm{ev}_\mu(d^2F_\alpha(\nu)). \end{aligned}$$

如果定义域空间 V 是笛卡儿空间 $\mathbb{R}^n$, 则 $(D_{\delta^j}F)(\cdot) = (\partial F/\partial x_j)(\cdot)$ 在 $\boldsymbol{a}$ 的可微性由定理 9.2 意味着二阶偏导数 $(\partial^2 F/\partial x_i \partial x_j)(\boldsymbol{a})$ 的存在性, 并且让 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 固定, 我们则有

$$\begin{aligned} D_{\boldsymbol{c}}(D_{\boldsymbol{b}}F) &= D_{\boldsymbol{c}}\left(\sum b_i \frac{\partial F}{\partial x_i}\right) = \sum b_i D_{\boldsymbol{c}} \frac{\partial F}{\partial x_i} \\ &= \sum b_i \left(\sum c_j \frac{\partial}{\partial x_j}\left(\frac{\partial F}{\partial x_i}\right)\right) = \sum_{i,j} b_i c_j \frac{\partial^2 F}{\partial x_j \partial x_i}. \end{aligned}$$

因此,

推论 16.1 如果在上面定理中 $V = \mathbb{R}^n$, 则 $d^2F_{\boldsymbol{a}}$ 的存在性意味着所有二阶偏导数 $(\partial^2 F/\partial x_i \partial x_j)(\boldsymbol{a})$ 的存在性并且

$$d^2F_{\boldsymbol{a}}(\boldsymbol{b}, \boldsymbol{c}) = D_{\boldsymbol{b}}(D_{\boldsymbol{c}}F)(\boldsymbol{a}) = \sum_{i,j=1}^{n} b_i c_j \frac{\partial^2 F}{\partial x_j \partial x_i}(\boldsymbol{a}).$$

另外, 从上面的思考以及定理 9.3 我们也有如下结论:

定理 16.2 如果 $V=\mathbb{R}^n$, 且如果所有二阶偏导数 $(\partial^2 F/\partial x_i \partial x_j)(\boldsymbol{a})$ 存在并在开集 A 上连续, 则二阶微分 $d^2F_{\boldsymbol{a}}$ 在 A 上存在且连续.

证明 从定理 9.3 我们直接有每个一阶偏导数 $(\partial F/\partial x_j(\cdot)$ 可微. 但 $\partial F/\partial x_j = \mathrm{ev}_{\delta j} \circ dF$, 那么此定理是下面一般原理的推论. □

引理 16.1 如果 $\{S_i\}_1^k$ 是在向量空间 W 上线性映射的有限组, 使得 $S = \prec S_1, \cdots, S_k \succ$ 可逆, 则映射 $F: A \to W$ 在 α 可微当且仅当 $S_i \circ F$ 在 α 对所有的 i 可微.

证明 由定理 8.1 和 6.2 知这是由于 $S \circ F$ 和 $F = S^{-1} \circ S \circ F$ 可微. □

这些考量明显地可推广到任意个数的微分. 因此, 如果 $d^2F_{(\cdot)}: y \mapsto d^2F_y$ 在 $\boldsymbol{a}$ 可微, 则对固定的 $\boldsymbol{b}$ 和 $\boldsymbol{c}$, 赋值 $d^2F_{(\cdot)}(\boldsymbol{b}, \boldsymbol{c})$ 在 $\boldsymbol{a}$ 可微, 并且公式

$$(D_{\boldsymbol{c}}(D_{\boldsymbol{b}}F))(\cdot) = d^2F_{(\cdot)}(\boldsymbol{b}, \boldsymbol{c}) = \sum_{i,j} b_i c_j \frac{\partial^2 F}{\partial x_j \partial x_i}(\cdot)$$

表明 (对特别选取的 $\boldsymbol{b}$ 和 $\boldsymbol{c}$) 所有的二阶偏导 $(\partial^2 F/\partial x_j \partial x_i)(\cdot)$ 在 $\boldsymbol{a}$ 可微, 并且有

$$(D_{\boldsymbol{d}} D_{\boldsymbol{c}} D_{\boldsymbol{b}} F)(\boldsymbol{a}) = D_{\boldsymbol{d}}((D_{\boldsymbol{c}} D_{\boldsymbol{b}} F)(\cdot))|_{\boldsymbol{a}} = \sum_{i,j,k} b_i c_j d_k \frac{\partial^3 F}{\partial x_k \partial x_j \partial x_i}(\boldsymbol{a}).$$

反之, 如果所有三阶偏导存在并在 A 上连续, 则由定理 9.3 知二阶偏导在 A 上可微, 于是由引理知 $d^2F_{(\cdot)}$ 可微, 这是因为 $(\partial^2 F/\partial x_i \partial x_j)(\cdot) = \mathrm{ev}_{\prec \delta^i, \delta^j \succ} \circ d^2F_{(\cdot)}$.

像读者记得的那样, 在处理高阶导数时 $\partial^2 F/\partial x_i \partial x_j = \partial^2 F/\partial x_j \partial x_i$ 具有极端重要性, 我们在这里也非常需要这个相同的定理.

定理 16.3 二阶微分是它两个变量的对称函数:

$$(d^2F_\alpha(\eta))(\xi) = (d^2F_\alpha(\xi))(\eta).$$

证明 由 $d(dF)_\alpha$ 的定义, 对给出的 ε, 存在 δ 使得当 $\|\eta\| < \delta$, $\|\Delta(dF)_\alpha(\eta) - d(dF)_\alpha(\eta)\| \leqslant \|\eta\|$. 自然, $\Delta(dF)_\alpha(\eta) = dF_{\alpha+\eta} - dF_\alpha$. 如果我们以 $\eta + \zeta$ 替换 η 写下同一不等式, 则这两个不等式左端中两个变换的差是

$$(dF_{\alpha+\eta} - dF_{\alpha+\eta+\zeta}) - d^2F_\alpha(-\zeta),$$

而三角不等式表明

$$\|(dF_{\alpha+\eta} - dF_{\alpha+\eta+\zeta}) - d^2F_\alpha(-\zeta)\| \leqslant \varepsilon(\|\eta\| + \|\eta+\zeta\|) \leqslant 2\varepsilon(\|\eta\| + \|\zeta\|),$$

其中假定了 η 和 $\eta+\xi$ 的范数最多为 δ. 我们取 $\|\zeta\| \leqslant 8/3$, $\|\eta\| \leqslant 2\delta/3$. 如果让 ζ 固定并令 $T = d^2F_\alpha(-\zeta)$ 和 $G(\xi) = F(\xi) - F(\xi+\zeta)$, 则此不等式成为 $\|dG_{\alpha+\eta} - T\| \leqslant 2\varepsilon(\|\eta\| + \|\zeta\|)$, 并且由于它只要 $\|\eta\| \leqslant 2\delta/3$ 就成立, 故我们可应用定理 7.4 的推论, 得到在假设 η 和 $\eta+\xi$ 的范数最多为 $2\delta/3$ 时有 $\|\Delta G_{\alpha+\eta}(\xi) - T(\xi)\| \leqslant 2\varepsilon(\|\eta\| + \|\zeta\|)\|\xi\|$ 的结果. 因此如果 η, ζ 和 ξ 的范数都是最多为 $\delta/3$ 时, 此不等式成立. 如果现在令 $\zeta = -\eta$, 我们有

$$\Delta G_{\alpha+\eta} = \Delta F_{\alpha+\eta} - \Delta F_{\alpha+\eta+\zeta} = \Delta F_{\alpha+\eta} - \Delta F_\alpha,$$

那么 $\Delta G_{\alpha+\eta}(\xi) = F(\alpha+\eta+\xi) - F(\alpha+\eta) - F(\alpha+\xi) + F(\alpha)$. 这个 η 和 ξ 的函数被称作 F 在 α 的*二阶差分*, 以 $\Delta^2F_\alpha(\eta,\xi)$ 表示. *注意它是对 ξ 和 η 对称的.* 我们最后的那个不等式现在可以重写为

$$\|\Delta^2F_\alpha(\eta,\xi) - (d^2F_\alpha(\eta))(\xi)\| \leqslant 4\varepsilon\|\eta\|\|\xi\|.$$

调换 η 和 ξ 并用 Δ^2F_α 的对称性, 我们看到

$$\|(d^2F_\alpha(\eta))(\xi) - (d^2F_\alpha(\xi))(\eta)\| \leqslant 8\varepsilon\|\eta\|\|\xi\|,$$

其中假定 η 和 ξ 的范数最多为 $\delta/3$. 现在由通常的齐性论证知道这个不等式对所有的 η 和 ξ 都成立. 最后, 因为 ε 是任意的, 故左端为 0. □

读者必会记得在初等微积分中，一个函数 f 的一个判别点 a[即 $f'(a) = 0$], 如果有 $f''(a)$ 存在且非零, 则它是一个相对极值点. 事实上, 如果 $f''(a) < 0$, 则 f 在 a 有个相对极大值, 这是因为 $f''(a) < 0$ 表明 f' 在 a 的一个邻域中递降, 因而 f 的图像在 a 的邻域中凹向下. 类似地, 如果 $f'(a) = 0$ 而 $f''(a) > 0$ 则 f 在 a 有一个相对极小值. 如果 $f''(a) = 0$ 则无定论.

如果 f 是定义于有限维向量空间 V 中开集 A 上的实函数, 又如果 $\alpha \in A$ 是 f 的一个判别点并且如果 d^2f_α 存在, 而且是 $\mathrm{Hom}(V, V^*)$ 中的一个非异元素, 我们则可以对于 f 在 α 附近的行为得出相似的结论, 差别仅在于现在有更为丰富的各种可能性. 读者大概已经对从 $\mathbb{R}^2$ 到 $\mathbb{R}$ 的函数会发生什么情况有些熟悉了. a 会是一个相对极大点 (f 图像上的一个 "帽" 点), 一个相对极小点, 或是一个像图 3.13 中显示的鞍点, 这是一个平移了的函数 Δf_a 的图像. 但是应该认识到

对于鞍形曲面的这个定位, 新的轴必须选成图中所显示的那样的轴. 用 Δf_a 代替 f 等于假定了 0 是判别点且 $f(0)=0$. 注意, 如果 0 是个鞍点, 则有两个余子空间, 即图 3.13 中的坐标轴, 使得当 f 限制于其中一条时, 0 为 f 的相对极大点, 而限制在另一条时为相对极小点.

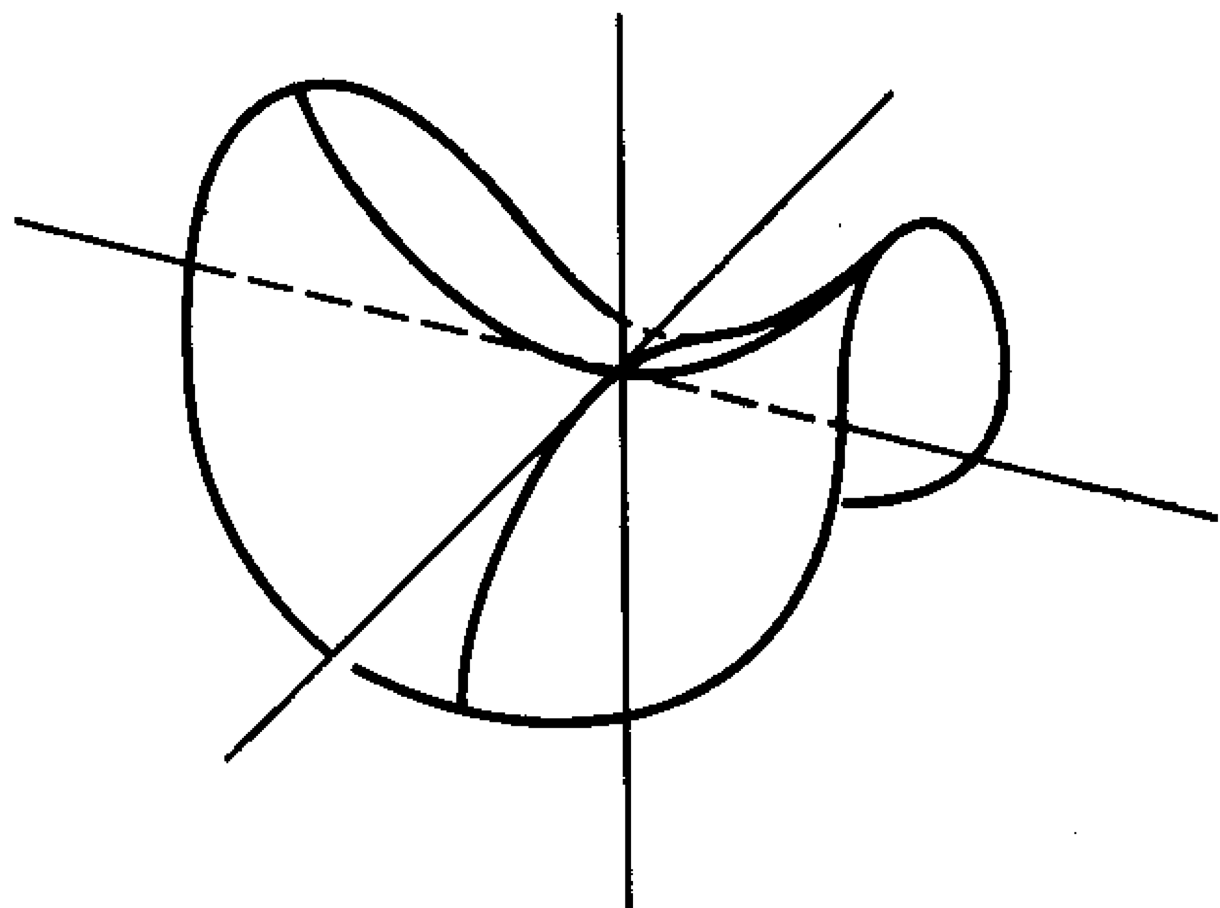


图 3.13

我们现在将研究一般的情形并发现它恰似二维的情形, 但要除去鞍点的情形, 这时使此判别点为极大点的子空间的维数可以是从 1 到 $n-1$ 中任一个数 (其中 $d(V)=n$). 另外, 此维数恰好是二次型 $q(\xi)=\omega(\xi,\xi)=d^2f_\alpha(\xi,\xi)$ 在标准法正交基表示下 -1 的个数.

那么我们的假设条件是 f 为有限维赋范线性空间 V 上一个开子集 A 上的连续可微的实函数, $\alpha\in A$ 是 f 的一个判别点 ($df_\alpha=0$), 并且映射 $d^2f_\alpha: V\to V^*$ 存在且非异. 最后的这个假设等价于假定双线性形式 $\omega(\xi,\eta)=d^2f_\alpha(\xi,\eta)$ 关于 V 中任一基有一个非异的矩阵. 我们现在用第二章定理 7.1 选取一个 ω 法正交基 $\{\alpha_i\}_1^n$. 回想这表明了如果 $i\neq j$, 则 $\omega(\alpha_i,\alpha_j)=0$, 但对 $i=1,\cdots,p,\omega(\alpha_i,\alpha_i)=1$, 时 $i=p+1,\cdots,n,\omega(\alpha_i,\alpha_i)=-1$. 因为矩阵 $t_{ij}=\omega(\alpha_i,\alpha_j)$ 非异, $\omega(\alpha_i,\alpha_i)$ 永不为零: 如果 $\omega(\alpha_i,\alpha_i)=0$ 则整个第 i 列都是 0, 那么列空间具维数 $\leqslant n-1$, 从而此矩阵为奇异.

我们可以用基同构 φ 将 $\mathbb{R}^n$ 替换掉 V(即用 $f\circ\varphi$ 替换 f), 故可假定 $V=\mathbb{R}^n$, 标准基为 ω 法正交, 具有 $\omega(\boldsymbol{x},\boldsymbol{y})=\sum_1^p x_iy_i-\sum_{p+1}^n x_iy_i$. 由于

$$\omega(\delta^i,\delta^j)=d^2f_{\boldsymbol{a}}(\delta^i,\delta^j)=D_{\delta^i}D_{\delta^j}f(\boldsymbol{a})=\frac{\partial^2F}{\partial x_i\partial x_j}(\boldsymbol{a}),$$

我们的 ω 法正交性假设就是对 $i\neq j$ 时 $(\partial^2f/\partial x_i\partial x_j)(\boldsymbol{a})=0$, 而对 $i=1,\cdots,p$, $\partial^2f/\partial x_i^2=1$, 对 $i=p+1,\cdots,n,\partial^2f/\partial x_i^2=-1$. 由于 p 可以有从 0 到 n 的任何

值, 故有 $n+1$ 种可能性. 我们首先证明如果 $p=n$, 则 $\boldsymbol{a}$ 是 f 的一个相对极小. 这时称二次型 q 为正定的, 这是因为 $q(\boldsymbol{x})=\omega(\boldsymbol{x},\boldsymbol{x})$ 对任意非零 $\boldsymbol{x}$ 是正的. 我们也说双线型 $\omega(\boldsymbol{x},\boldsymbol{y})=d^2f_{\boldsymbol{a}}(\boldsymbol{x},\boldsymbol{y})$ 是正定的, 而以第五章的话说, 即 ω 是个内积 (纯量积).

定理 16.4 设 f 是定义在 $\mathbb{R}^n$ 中开子集 A 上的连续可微实函数, 设 $\boldsymbol{a}\in A$ 是 f 的一个判别点, 在此点 d^2f 存在且正定. 则 f 在 $\boldsymbol{a}$ 有一个相对极小值.

证明 像前面一样, 我们假设标准基 $\{\delta^i\}_1^n$ 是 ω 法正交的. 由 $d^2f_{\boldsymbol{a}}$ 的定义, 对给出的 ε, 存在 δ 使得只要 $\|\boldsymbol{y}\|\leqslant\delta$ 便有

$$|(df_{\boldsymbol{a}+\boldsymbol{y}}(\boldsymbol{x})-df_{\boldsymbol{a}}(\boldsymbol{x}))-d^2f_{\boldsymbol{a}}(\boldsymbol{x},\boldsymbol{y})|\leqslant\varepsilon\|\boldsymbol{x}\|\ \|\boldsymbol{y}\|.$$

现有 $df_{\boldsymbol{a}}=0$, 这是由于 $\boldsymbol{a}$ 是 f 的一个判别点, 另外 $d^2f_{\boldsymbol{a}}(\boldsymbol{x},\boldsymbol{y})=\sum_1^n x_iy_i$, 这是因为 $\{\delta^i\}_1^n$ 为 ω 法正交. 因此, 如果我们在 $\mathbb{R}$ 上使用 2 - 范数, 并令 $\boldsymbol{y}=t\boldsymbol{x}$, 我们则有

$$|df_{\boldsymbol{a}+t\boldsymbol{x}}(\boldsymbol{x})-t\|\boldsymbol{x}\|^2|\leqslant\varepsilon t\|\boldsymbol{x}\|^2.$$

又, 如果 $h(t)=f(\boldsymbol{a}+t\boldsymbol{x})$ 则 $h'(s)=df_{\boldsymbol{a}+s\boldsymbol{x}}(\boldsymbol{x})$, 而这个不等式便说的是 $(1-\varepsilon)t\|\boldsymbol{x}\|^2\leqslant h'(t)\leqslant(1+\varepsilon)t\|\boldsymbol{x}\|^2$. 积分并记住 $h(1)-h(0)=f(\boldsymbol{a}+\boldsymbol{x})-f(\boldsymbol{a})=\Delta f_{\boldsymbol{a}}(\boldsymbol{x})$, 我们有当 $\|x\|\leqslant\delta$, 则

$$\left(\frac{1-\varepsilon}{2}\right)\|\boldsymbol{x}\|^2\leqslant\Delta f_{\boldsymbol{a}}(\boldsymbol{x})\leqslant\left(\frac{1+\varepsilon}{2}\right)\|\boldsymbol{x}\|^2.$$

这证明不但 $\boldsymbol{a}$ 是个相对极小点而且只要 x 充分地小, $\Delta f_{\boldsymbol{a}}$ 就在两个非常靠近的抛物体之间. □

上面的论证对一般情形同样可行得通. 如果

$$q(\boldsymbol{x})=\sum_1^p x_i^2-\sum_{p+1}^n x_i^2$$

为二阶微分的二次型, 且 $\|\boldsymbol{x}\|_2^2=\sum_1^n x_i^2$, 则在上面不等式中换去绝对值内的 $\|\boldsymbol{x}\|^2$ 为 $q(\boldsymbol{x})$, 我们得到现在的

$$\frac{q(\boldsymbol{x})-\varepsilon\|\boldsymbol{x}\|^2}{2}\leqslant\Delta f_{\boldsymbol{a}}(\boldsymbol{x})\leqslant\frac{q(\boldsymbol{x})+\varepsilon\|\boldsymbol{x}\|^2}{2},$$

或者说

$$\frac{1}{2}\left(\sum_{1}^{p}(1-\varepsilon)x_i^2-\sum_{p+1}^{n}(1+\varepsilon)x_i^2\right)\leqslant \Delta f_{\boldsymbol{a}}(\boldsymbol{x})$$
$$\leqslant \frac{1}{2}\left(\sum_{1}^{p}(1+\varepsilon)x_i^2-\sum_{p+1}^{n}(1-\varepsilon)x_i^2\right).$$

这表明 $\Delta f_{\boldsymbol{a}}$ 当 $\|x\|<\delta$ 时位于两个相同类型的非常靠近的二次曲面之间. 如果 $1\leqslant p\leqslant n-1$ 和 $\boldsymbol{a}=0$, 则 f 在子空间 $V_1=L(\{\delta^i\}_1^p)$ 上有一个相对极小, 而在补空间 $V_2=L(\{\delta^i\}_{p+1}^n)$ 上有个相对极大.

按照我们在 2.7 节末的评注, 我们能写全一个二元变量函数的判别点类型而没有进行法正交化, 这只要观察 (非异的) 形式 $d^2f_{\boldsymbol{a}}$ 的矩阵的行列式就可以了. 此行列式是

$$t_{11}t_{22}-(t_{12})^2=\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2}-\left(\frac{\partial^2 f}{\partial x_1\partial x_2}\right)^2.$$

如果它是正的, 则 $\boldsymbol{a}$ 或是相对极小或相对极大. 我们可以说出到底是哪一个, 这只要沿着一条单个直线, 譬如 x_1 轴去观察 f. 因此如果 $\frac{\partial^2 f}{\partial x_1^2}<0$, 那么 $\boldsymbol{a}$ 是个相对极大点. 另一方面, 如果上面的表达式为负, $\boldsymbol{a}$ 则是个鞍点.

对变分法重要之处是当定义域空间被换为我们将在下一章研究的一般类型的空间时能保持定理 16.4 不变; 这个一般性的空间被称为巴拿赫空间. 这时假设条件是 α 为 f 的一个判别点, $q(\xi)=d^2f_\alpha(\xi,\xi)$ 正定, 以及内积范数 $q^{1/2}$(见第五章) 等价于 V 上给定的范数. 此证明真正地保持了不变.

*3.17 高阶微分; 泰勒公式

我们已经看到如果 V 和 W 为赋范线性空间, 并且如果 F 为从 V 中开集 A 到 W 的可微映射, 则它的微分 $dF=dF_{(\cdot)}$ 是从 A 到 $\mathrm{Hom}(V,W)$ 的映射. 如果这个映射在 A 上可微, 则它的微分 $d(dF)=d(dF)_{(\cdot)}$ 是从 A 到 $\mathrm{Hom}(V,\mathrm{Hom}(V,W))$ 的一个映射. 我们记得, $\mathrm{Hom}(V,\mathrm{Hom}(V,W))$ 中一个元素由对偶性等价于从 $V\times W$ 到 W 的双线性映射, 而如果我们将所有这些双线性映射的空间以 $\mathrm{Hom}^2(V,W)$ 表示, 则 $d(dF)$ 可以考虑成从 A 到 $\mathrm{Hom}^2(V,W)$ 的映射. 我们写成 $d(dF)=d^2F$, 并称此映射为 F 的二阶微分. 在 16 节中我们看到

过 $d^2F_\alpha(\xi,\eta)=D_\xi(D_\eta F)(\alpha)$, 并且如果 $V=\mathbb{R}^n$, 则

$$d^2F_{\boldsymbol{a}}(\boldsymbol{b},\boldsymbol{c})=D_{\boldsymbol{b}}D_{\boldsymbol{c}}F(\boldsymbol{a})=\sum b_ic_j\frac{\partial^2 F}{\partial x_j\partial x_i}(\boldsymbol{a}).$$

高阶微分以相同的方式定义. 如果

$$d^2F:A\to \mathrm{Hom}^2(V,W)$$

在 A 上可微, 则它的微分 $d(d^2F)=d^3F$ 是从 A 到 $\mathrm{Hom}(V,\mathrm{Hom}^2(V,W))=\mathrm{Hom}^3(V,W)$, 这是从 $V^3=V\times V\times V$ 到 W 的所有三重线性映射的空间. 归纳地继续下去, 我们得到了 F 在 A 上的等 n 阶微分的概念, 它是从 A 到 $\mathrm{Hom}(V,\mathrm{Hom}^{n-1}(V,W))=\mathrm{Hom}^n(V,W)$ 的一个映射, 后者是从 V^n 到 W 的所有的 n 重线性映射. d^2F_α 是 $\mathrm{Hom}^2(V,W)$ 中对称元素的定理可归纳地推广去证明 d^nF_α 是 $\mathrm{Hom}^n(V,W)$ 的一个对称元. 我们略去此证明.

关于二阶导数用混合方向导数赋值的定理也可归纳地推广, 给出了

$$D_{\xi_1},\cdots,D_{\xi_n}F(\alpha)=d^nF_\alpha(\xi_1,\cdots,\xi_n),$$

因为从左边的项开始, 我们有

$$\begin{aligned}
D_{\xi_1}(D_{\xi_2},\cdots,D_{\xi_n}F)(\cdot)|_\alpha &= d(D_{\xi_2},\cdots,D_{\xi_n}F(\cdot))_\alpha(\xi_1)\\
&= d(d^{n-1}F_{(\cdot)}(\xi_2,\cdots,\xi_n))_\alpha(\xi_1)\\
&= d(\mathrm{ev}_{\prec\xi_2,\cdots,\xi_n\succ}\circ d^{n-1}F_{(\cdot)})_\alpha(\xi_1)\\
&= [\mathrm{ev}_{\prec\xi_2,\cdots,\xi_n\succ}\circ d(d^{n-1}F_{(\cdot)})_\alpha](\xi_1)\\
&= \mathrm{ev}_{\prec\xi_2,\cdots,\xi_n\succ}(d^nF_\alpha(\xi_1))\\
&= (d^nF_\alpha(\xi_1))(\xi_2,\cdots,\xi_n)\\
&= d^nF_\alpha(\xi_1,\cdots,\xi_n).
\end{aligned}$$

如果 $V=\mathbb{R}^n$, 则关于偏导数的结论归纳地推广到按相同方式证明 F 在 A 上有直到 m 阶连续微分当且仅当所有 m 阶偏导数 $\partial^mF/\partial x_{i_1},\cdots,\partial x_{i_m}$ 存在并在 A 上连续, 并有

$$d^mF_{\boldsymbol{a}}(c^1,\cdots,c^n)=\sum_{i_1,\cdots,i_m=1}^{n}c^{i_1},\cdots,c^{i_m}\frac{\partial^m F}{\partial x_{i_1}\cdots\partial x_{i_m}}(\boldsymbol{a}).$$

我们现在考虑 F 沿直线 $t \mapsto \alpha + t\eta$ 的行为, 自然这里的 α 和 η 是固定的. 如果 $\lambda(t) = F(\alpha + t\eta)$, 我们则可归纳地证明

$$d^j\lambda/dt^j = (D_\eta)^j F(\alpha + t\eta).$$

由定理 7.2 我们知道这对 $j = 1$ 是对的, 现假定它对 $j = m$ 为真, 那么由同一定理我们有

$$\begin{aligned}\frac{d^{m+1}\lambda}{dt^{m+1}} = \left(\frac{d^m\lambda}{dt^m}\right)'(t) = d(D_\eta^m F)_{\alpha+t\eta}(\eta) = D_\eta(D_\eta^m)(\alpha + t\eta)\\ = D_\eta^{m+1}F(\alpha + t\eta).\end{aligned}$$

现在假设 F 是实的 $(W = \mathbb{R})$. 我们则有泰勒公式:

$$\lambda(t) = \lambda(0) + t\lambda'(0) + \cdots + \frac{t^m}{m!}\lambda^m(0) + \frac{t^{m+1}}{(m+1)!}\lambda^{(m+1)}(kt),$$

其中 k 在 0 和 1 之间, 取 $t = 1$ 并用上面的进行替换, 我们有

$$F(\alpha + \eta) = F(\alpha) + D_\eta F(\alpha) + \cdots + \frac{1}{m!}D_\eta^m F(\alpha) + \frac{1}{(m+1)!}D_\eta^{m+1}F(\alpha + k\eta),$$

它是在赋范线性空间背景下的一般泰勒公式. 用微分表示, 它是

$$\begin{aligned}F(\alpha + \eta) = F(\alpha) + dF_\alpha(\eta) + \cdots + \frac{1}{m!}d^m F_\alpha(\eta, \cdots, \eta)\\ + \frac{1}{(m+1)!}d^{m+1}F_{\alpha+k\eta}(\eta, \cdots, \eta).\end{aligned}$$

如果 $V = \mathbb{R}^n$, 则 $D_{\boldsymbol{y}}G = \sum_1^n y_i \partial G/\partial x_i$, 故泰勒展式中的通项为

$$\frac{1}{m!}\left(\sum_1^n y_i \frac{\partial}{\partial x_i}\right)^m F(\boldsymbol{a}) = \frac{1}{m!}\sum_{i_1, \cdots, i_m = 1}^n y_{i_1}\cdots y_{i_m}\frac{\partial^m F}{\partial x_{i_1}\cdots\partial x_{i_m}}(\boldsymbol{a}).$$

如果 $m = n = 2$, 并且如用记号 $\boldsymbol{x} = \prec x, y \succ$, $\boldsymbol{s} = \prec s, t \succ$, 则

$$\frac{1}{2!}D_{\boldsymbol{s}}^2 F(\boldsymbol{a}) = \frac{1}{2}\left[s^2\frac{\partial^2 F}{\partial x^2}(\boldsymbol{a}) + 2st\frac{\partial^2 F}{\partial x \partial y}(\boldsymbol{a}) + t^2\frac{\partial^2 F}{\partial y^2}(\boldsymbol{a})\right].$$

上面的描述在逻辑上是简单的, 但它效率不高, 因为它重复了相同的项; 譬如像 $y_1 y_2(\partial^2 F/\partial x_1 \partial x_2)$ 和 $y_2 y_1(\partial^2 F/\partial x_2 \partial x_1)$. 我们最后对感兴趣的读者描述一下对此极繁复的情形的现代 "多重指标" 的记号.

记住我们在考虑的是 F 的泰勒公式的第 m 项, 而 F 有 n 个变量.

对任意的 n 联组 $\boldsymbol{k} = \prec k_1, \cdots, k_n \succ$, 其中 k_i 为非负整数, 我们定义 $|\boldsymbol{k}|$ 为 $\sum_1^n k_i$, 并对 $x \in \mathbb{R}^n$, 我们令 $\boldsymbol{x}^{\boldsymbol{k}} = x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$. 我们也令 $F_{\boldsymbol{k}} = F_{k_1 \cdots k_n}$, 或者更好地, 如果 $D_j F = \partial F / \partial x_j$, 我们令

$$D^{\boldsymbol{k}} F = D_1^{k_1} D_2^{k_2} \cdots D_n^{k_n} F = F_{\boldsymbol{k}}.$$

最后, 令 $\boldsymbol{k}! = k_1! k_2! \cdots k_n!$, 并且如果 $p \geqslant |\boldsymbol{k}|$, 我们令 $\binom{p}{\boldsymbol{k}} = p! / \boldsymbol{k}!(p - |\boldsymbol{k}|)!$. 于是 F 的泰勒展式的第 m 项为

$$\frac{1}{m!} \sum_{|\boldsymbol{k}| = m} \binom{m}{\boldsymbol{k}} D^{\boldsymbol{k}} F(\boldsymbol{a}) \boldsymbol{x}^{\boldsymbol{k}},$$

这的确是属于记号法的胜利.

对于能广泛应用而言, 一般的泰勒公式显得过于繁琐了; 它主要是具有理论上的价值. 我们实际计算泰勒展开式是用其他手段求出来的, 如像将一个多项式 (或幂级数) 代入一个幂级数. 例如,

$$\begin{aligned} \sin(x + y^2) &= (x + y^2) - \frac{(x + y^2)^3}{3!} + \frac{(x + y^2)^5}{5!} \cdots \\ &= x + y^2 - \frac{x^3}{3!} - \frac{x^2 y^2}{2} + \left(\frac{x^5}{5!} - \frac{x y^4}{2} \right) \\ &\quad + \left(\frac{x^4 y^2}{4!} - \frac{y^6}{3!} \right) \cdots . \end{aligned}$$

一个从 A 到 W 的具有一直到 k 的所有阶的连续微分的映射被称作是 A 上的类 C^k 中的元. 所有这种映射的汇集被记为 $C^k(A, W)$ 或 $\mathfrak{C}^k(A, W)$, 显见, $C^k(A, W)$ 是个向量空间 (归纳). 另外, 可以归纳证明 C^k 映射的复合仍然属于类 C^k 本身. 这有赖于了解到复合 $F \circ G$ 的 m 阶微分的一般形式是个有限和, 而它的每项是由 $F, dF, \cdots, d^m F, G, dG, \cdots, d^m G$ 中选出的一些函数的复合.

在这些计算中涉及到多元函数, 对每个这种函数的处理最简单的是当作一个单一的 n 联组变量的函数并应用定理 8.1 的明显推论, 即如果 $G^1, \cdots, G^n$ 属类 C^k, 则 $G = \prec G^1, \cdots, G^n \succ$ 也是, 且 $d^k G = \prec d^k G^1, \cdots, d^k G^n \succ$. 作为复合的特殊情形, 我们得出结论说, C^k 映射的积仍是 C^k 类的.

我们将在下一章看到, $\varphi: T \mapsto T^{-1}$ 是 $\mathrm{Hom} V$(如果 V 为巴拿赫空间) 中可逆元素组成的开集上的可微映射, 并且 $d_{\varphi(T)}(H) = -T^{-1} H T^{-1}$. 由于 $\prec S, H, T \succ \mapsto S^{-1} H T^{-1}$ 因而具有了连续的偏微分, 我们可以继续, 并用另一个

归纳法证明了 φ 对每个 k 属类 C^k, 并且 $d^m_{\varphi(T)}(H_1, \cdots, H_m)$ 是 $T^{-1}, H_1, \cdots, H_m$ 的微分的有限积的有限和. 于是可以得出, 为一个 C^k 函数 G 隐式定义的函数 F 仍是类 C^k 的, 这是因为它的微分像在隐函数定理中的计算那样, 是那些类 C^{k-1} 的映射的复合.

称一个对所有 k 属于类 C^k 的映射 F 为属于类 C^∞ 的, 并且从我们上面的评注得出 C^∞ 映射族在所有我们在微积分中遇到过的运算下是封闭的结论. 如果 F 的定义域是 $\mathbb{R}^n$ 中的开集 A, 则 $F \in \mathfrak{C}^\infty(A, W)$ 当且仅当 F 所有的偏导数存在并在 A 上连续.

第四章　紧性和完备性

在这一章里我们要研究一个赋范线性空间V 的子集的两个性质，它们与这样的事情有关：即按某种意义说所有应该存在的点确实存在. 这些概念大体上与 V 的代数结构无关，因而我们将在它们自己的最自然的架构即*度量空间*之下研究它们. 这两个性质中较强的那个*紧性*有助于向我们解释为什么有限维空间的理论是如此简单和令人满意的. 那个较弱的性质即*完备性*是与一些重要的无穷维赋范线性空间所共享的性质，它让我们能以几乎满意的方式来处理这些空间.

正是这些性质把微积分从主要是一种形式的理论中拯救出来. 它们允许我们以极限过程去定义各种至关重要的要素，例如可归功于它们的有无穷级数具有和，连续实函数达到极大值以及定积分的*存在*. 对于实数系自身，紧性等价于上确界性质，在第三章我们的微学理论构造中它已经是一个绝对实质性的工具了.

在第 8 节到第 10 节中我们要把完备性应用到微积分上. 这几节中最前面专门讲了由幂级数定义的函数的存在性和可微性，并且由于我们要把幂级数算入一个算子 T 中，我们将趁此机会引进并展开巴拿赫代数的概念. 然后我们将证明压缩映射不动点定理，它是我们在第三章中隐函数定理的未完成证明的缺失成分，而它还是第六章的常微分方程的基本存在性和惟一性定理的基础. 第 10 节中我们将证明将线性映射拓展到一个完备赋范线性空间的简单定理并用它去构造参数弧的黎曼积分.

4.1 度量空间; 开集和闭集

在前一章我们偶尔处理过在定义域为一赋范线性空间 V 中任一子集 A 的情形中的收敛性和连续性的问题. 在这样的讨论中, V 的代数结构衰减为一个背景, V 的向量运算仅仅用来产生一个组合 $\|\alpha-\beta\|$, 它被解释为从 α 到 β 的距离. 如果我们从那些对收敛性和连续性论证的本质性的各种情况中进行提炼, 我们则发现我们需要一个空间 A 和一个函数 $\rho: A\times A\to\mathbb{R}$, 称 $\rho(x,y)$ 为从 x 到 y 的*距离*,使得

(1) $\rho(x,y)>0$, 如果 $x\neq y$, 而 $\rho(x,x)=0$;

(2) $\rho(x,y)=\rho(y,x)$ 对所有 $x,y\in A$ 成立.

(3) $\rho(x,z)\leqslant\rho(x,y)+\rho(y,z)$ 对所有 $x,y,z\in A$ 成立.

任何一个集合 A 连同一个从 $A\times A$ 到 $\mathbb{R}$ 的这样一个函数 ρ 被称作一个*度量空间*; 函数 ρ 是其*度量*. 显然赋范线性空间在范数度量 $\rho(\alpha,\beta)=\|\alpha-\beta\|$ 下是个度量空间, 一个度量空间的任意子集 B 在 $\rho\restriction B\times B$ 下自身也是个度量空间. 如果我们从一个具有好的直观的空间, 譬如像在其中一个标准范数下的 $\mathbb{R}^n$ 开始着手, 选取一个怪模怪样的子集 B, 显见那样一个度量空间是个非常奇怪的东西, 可能不会有差不多我们能想到的任何性质.

实际上度量空间经常作为在范数度量下的赋范线性空间的子集出现,但它们也有其他的来源. 甚至在赋范线性空间的范围中, 也会使用不同于范数度量的其他度量. 举例来说, S 是 $\mathbb{R}^3$ 中一个二维球面, 譬如 $S=\{\boldsymbol{x}:\sum_1^3 x_i^2=1\}$, $\rho(\boldsymbol{x},\boldsymbol{y})$ 为从 $\boldsymbol{x}$ 到 $\boldsymbol{y}$ 的大圆距离. 或者, 更一般地, S 是 $\mathbb{R}^3$ 中任意一个光滑的二维曲面, 而 $\rho(\boldsymbol{x},\boldsymbol{y})$ 是联结 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 在 S 中的最短曲线的长度.

在本章我们只要合适就对论证采用度量空间的背景, 所以同学们要对此更一般却非常直观的概念变得熟悉起来. 我们从以度量的语言重新创造基本定义开始. 因为纯量向量的二分法在这个背景下不再是个因素, 我们将抛弃我们的约定, 即点用希腊字母表示或用黑体罗马字母表示, 而不管在哪里均使用我们想使用的字母.

定义 1.1 *如果 X 和 Y 是度量空间, 则 $f:X\to Y$ 在 $a\in X$ 连续是说如果对每个 ε, 存在一个 δ 使得*

$$\rho(x,a)<\delta\ \Rightarrow\ \rho(f(x),f(a))<\varepsilon.$$

这里我对不同的空间度量使用了相同的记号 'ρ', 正如同早先对范数符号的模糊用法一样.

定义 1.2 在 p 的半径 r 的 (开) 球 $B_r(p)$, 简单地说, 是从 p 到其距离小于 r 的点的集合:

$$B_r(p) = \{x : \rho(x,p) < r\}.$$

定义 1.3 一个子集 $A \subset X$ 为开是说 A 中每个点都是某个包含在 A 中的球的中心, 即如果

$$(\forall p \in A)(\exists r^{>0})(B_r(p) \subset A).$$

引理 1.1 每个球均为开; 事实上, 如果 $q \in B_r(p)$, $\delta = r - \rho(p,q)$, 则 $B_\delta(q) \subset B_r(p)$.

证明 这等于三角不等式. 因为, 如果 $x \in B_\delta(q)$, 则 $p(x,q) < \delta$, 并且 $\rho(x,p) \leqslant \rho(x,q) + \rho(q,p) < \delta + \rho(p,q) = r$, 故而 $x \in B_r(p)$. 因此 $B_\delta(q) \subset B_r(p)$. □

引理 1.2 如果取 p 固定, 则 $\rho(p,x)$ 是 x 的连续函数.

证明 第三章引理 3.1 的一个符号对符号的释意表明 $|\rho(p,x) - \rho(p,y)| \leqslant \rho(x,y)$, 故 $\rho(p,x)$ 实际上是个具常数 1 的利普希茨函数. □

定理 1.1 度量空间 S 的所有开子集的族 $\mathfrak{F}$ 具有下列性质:

(1) 任意开集的汇集的并为开; 即如果 $\{A_i : i \in I\} \subset \mathfrak{F}$, 则 $\cup_{i\in I} A_i \in \mathfrak{F}$.

(2) 两个开集的交为开; 即如果 $A, B, \in \mathfrak{F}$ 则 $A \cap B \in \mathfrak{F}$.

(3) $\varnothing, V \in \mathfrak{F}$.

证明 这些性质可立刻由定义得到. 因此任意在 $\cup_i A_i$ 的点 p 必在某个 A_j 中, 而且因为 A_j 为开, p 的某个球因此必是 A_j 的一个子集, 从而也是更大的集合 $\cup_i A_i$ 的子集. □

推论 1.1 一个集合为开当且仅当它是开球的并.

证明 由开集的定义、上面的引理、定理的性质 (1) 得出. □

任意一个集合 A 的所有开子集的并由 (1) 是 A 的一个开子集, 因而是 A 的最大的开子集. 称它为 A 的内部,以 A^{int} 表示. 显见, p 在 A^{int} 中当且仅当 p 的某个球是 A 的子集, 把 A^{int} 形象地作为所有在 A 中球的并是有益的.

定义 1.4 如果 A' 为开则称 A 是闭的.

上面的定理和德摩根法则 (0.11 节) 于是产生了对于闭集合的性质的互补集合.

定理 1.2 (1) 任一族闭集的交集是闭的.

(2) 两个闭集的并为闭集.

(3) $\varnothing$ 和 V 都是闭的.

证明 例如, 假定 $\{B_i : i \in I\}$ 是一族闭集合. 则对每个 i, 补集 B_i' 为开, 故 U_iB_i' 由前面的定理为开. 又由德摩根法则(见 0.11 节), $\cup_i B_i' = (\cap_i B_i)'$. 因此 $\cap_i B_i$ 是一个开集的补集因而为闭. □

继续我们 “互补式” 的阐述. 我们定义一个任意集 A 的闭包 $\overline{A}$ 为包含 A 的所有闭集的交集, 于是由上面的 (1) 知, $\overline{A}$ 是包含 A 的最小闭集. 德摩根法则蕴涵了重要的等式

$$(\overline{A})' = (A')^{\text{int}}.$$

因为 F 是 A 的一个闭的超集当且仅当它的补集 $U = F'$ 是 A' 的一个开子集. 由德摩根法则, 所有这些集合 F 的交集的补集是所有这些集合 U 的并集. 那就是说, $\overline{A}$ 的补集是 $(A')^{\text{int}}$.

此等式产生了闭包的直接的特征描述:

引理 1.3 一个点 p 在 $\overline{A}$ 中当且仅当在 p 的每个球都与 A 相交.

证明 点 p 不在 $\overline{A}$ 中当且仅当 p 是 A' 的内点, 就是说, 当且仅当在 p 的某个球不与 A 相交. 这个等价关系最外面的两端的否定给出了引理. □

定义 1.5 任意子集 A 的边缘 ∂A 是其闭包与其内部的差集合. 因此有

$$\partial A = \overline{A} - A^{\text{int}}.$$

因为 $A - B = A \cap B'$, 我们有对称的特征描述 $\partial A = \overline{A} \cap \overline{(A')}$. 从而 $\partial A = \partial(A')$; 又,

$p \in \partial A$ 当且仅当在 p 的球既与 A 相交也与 A' 相交.

例题 1.1 球 $B_r(\alpha)$ 是开集. 在赋范线性空间中 $B_r(\alpha)$ 的闭包是在 α 的半径 r 的闭球 $\{\xi : \rho(\xi, \alpha) \leqslant r\}$. 它可从引理 1.3 容易看出. 于是边缘 $\partial B_r(a)$ 是在 α 的半径 r 的球面$\{\xi : \rho(\xi, \alpha) = r\}$. 如果把这曲面上某些但不是全部的点添加到

开球上, 我们便得到一个既非开也非闭的集合. 同学们应该预料到他所遇到的一个随意的集合会是个既非开又非闭的集合.

由下面的引理知连续函数提供了开集和闭集的重要来源.

引理 1.4 如果 X 和 Y 为度量空间, f 是从 X 到 Y 的连续映射, 则只要 A 是 Y 中开集, $f^{-1}[A]$ 便是 X 中开集.

证明 如果 $p \in f^{-1}[A]$, 则 $f(p) \in A$, 并且因为 A 为开, 某个球 $B_\varepsilon(f(p))$ 便是 A 的一个子集. 但是 f 在 p 点的连续正好是说, 存在一个 δ 使得 $f[B_\delta(p)] \subset B_\varepsilon(f(p))$. 特别, $f[B_\delta(p)] \subset A$ 从而 $B_\delta(p) \subset f^{-1}[A]$. 因此对每个 $p \in f^{-1}[A]$, 有一个关于 p 的球被含在 $f^{-1}[A]$ 中, 这就是说此集合为开. □

由于 $f^{-1}[A'] = (f^{-1}[A])'$, 我们也有下面的推论.

推论 1.2 如果 $X \to Y$ 连续, 则只要 C 是 Y 中闭集, $f^{-1}[C]$ 便是 X 中闭集.

把这两个结果的顺序前后颠倒也成立. 作为运用此引理的例题, 考虑对一固定的 $\alpha \in X$ 由 $f(\xi) = \rho(\xi, \alpha)$ 定义的连续函数 $f : X \to \mathbb{R}$. 集合 $(-r, r)$, $[0, r]$ 和 $\{r\}$ 分别为 $\mathbb{R}$ 的开, 闭和闭的子集. 因此它们在 f 下的逆像, 即球 $B_r(\alpha)$, 闭球 $\{\xi : \rho(\xi, \alpha) \leqslant r\}$ 和球面 $\{\xi : \rho(\xi, \alpha) = r\}$ 分别为 X 中的开, 闭和用的集合. 特别直接证明 $B_r(\alpha)$ 为开集的三角不等式的论证方式现在被看出已没有必要再借助于三角不等式论证以证明距离函数的连续性了 (引理 1.2).

但是在向前的方向上, 说连续函数把闭集带到闭集是不对的. 例如, 如果 $f : \mathbb{R} \to \mathbb{R}$ 是反正切函数, 则 $f[\mathbb{R}] = f$ 的值域 $= (-\pi/2, \pi/2)$, 它不是一个 $\mathbb{R}$ 的闭子集. 读者可能会感觉这个例子有蒙骗之处, 因为我们只预计一个闭集的 f - 像应在 f 值域这个度量空间中是个闭子集. 那他可以考虑从 $\mathbb{R}$ 到其值域 $[-1, 1]$ 的 $f(x) = 2x/(1 + x^2)$. 正整数集 $\mathbb{Z}^+$ 是 $\mathbb{R}$ 的闭子集, 但 $f[\mathbb{Z}^+] = \{2n/(1 + n^2)\}_1^\infty$, 在 $[-1, 1]$ 中不闭, 因为 0 显然在其闭包中.

定义两个非空集 A 和 B 之间距离 $\rho(A, B)$ 为 $\text{glb}\{\rho(a, b) : a \in A$ 和 $b \in B\}$. 如果 A 和 B 相交则此距离为 0. 如果 A 和 B 不交, 此距离仍然可能为 0. 例如, 平面中一个圆的内部和外部为不交的开集,它们相距为 0. x 轴和函数 $f(x) = 1/x$(的图像) 为不交的闭集,它们相距也为 0. 像我们早先曾评注的, 一个集合为闭当且仅当不在 A 中的每个点距 A 一个正的距离. 更一般地, 对任何集合 A, 一点 p 在

$\overline{A}$ 中当且仅当 $\rho(p, A) = 0$.

我们在下面列出一个赋范线性空间中两个子集间距离的性质.

(1) 在平移下距离不变:$\rho(A, B) = \rho(A+\gamma, B+\gamma)$
(因为 $\|(\alpha+\gamma)-(\beta+\gamma)\| = \|\alpha-\beta\|$).

(2) $\rho(kA, kB) = |k|\rho(A, B)$(因为 $\|k\alpha - k\beta\| = |k|\ \|\alpha-\beta\|$).

(3) 如果 N 是个子空间, 当通过 N 中一个向量平移 B 时则从 B 到 N 的距离不变:$\rho(N, B) = \rho(N, B+\eta)$, 如果 $\eta \in N$ (因为 $N+\eta = N$).

(4) 如果 $T \in \mathrm{Hom}(V, W)$, 则 $\rho(T[A], T[B]) \leqslant \|T\|\rho(A, B)$.
(因为 $\|T(\alpha) - T(\beta)\| \leqslant \|T\|\ \|\alpha-\beta\|$).

引理 1.5 *如果 N 是个真闭子空间, $0 < \varepsilon < 1$, 则存在 α 使得 $\|\alpha\| = 1$ 且 $\rho(\alpha, N) > 1-\varepsilon$.*

证明 选取任一 $\beta \notin N$. 于是 $\rho(\beta, N) > 0$(因为 N 为闭), 从而存在一个 $\eta \in N$, 使得

$$\|\beta - \eta\| < \rho(\beta, N)/1-\varepsilon$$

[由 $\rho(\beta, N)$ 的定义]. 令 $\alpha = (\beta-\eta)/\|\beta-\eta\|$. 则 $\|\alpha\| = 1$ 而

$$\begin{aligned}\rho(\alpha, N) &= \rho(\beta-\eta, N)/\|\beta-\eta\| = \rho(\beta, N)/\|\beta-\eta\| \\ &> \rho(\beta, N)(1-\varepsilon)/\rho(\beta, N) = 1-\varepsilon,\end{aligned}$$

这由 (2) 和 (3) 以及 η 的定义得到. □

读者可能会感到我们应该能够改进这个引理. 的确, 我们所必须做的一切是选择 N 中最靠近 β 的点, 故可得到 $\|\beta-\eta\| = \rho(\beta, N)$, 它最终给出一个向量 α 使 $\|\alpha\| = 1$ 而 $\rho(\alpha, N) = 1$. 但是这是我们的直观感觉让我们看走了眼: 如果 N 是无穷维的, 就可能不会有一个最近的点 η! 例如, 如同我们将在第五章的习题看到的, 如果 V 是空间 $\mathfrak{C}([-1, 1])$, 范数为 2 范数 $\|f\| = (\int_{-1}^{1} f^2)^{1/2}$, 并且如果 N 是 V 中那样的函数 g 的集合, 其中 $\int_0^1 g = 0$, 则 N 是个闭子空间, 对它我们不能找到这样一个 "最好" 的点 α. 但是如果 N 是有限维的, 我们总能找到这样一个点, 而且如果 V 是个希尔伯特空间 (见第五章) 我们也能找到.

习题

1.1 写出引理 1.2 的证明.

1.2 证明定理 1.1 的 (2) 和 (3).

1.3 证明定理 1.2 的 (2).

1.4 说一系列开集的交仍是开集, 这是不对的. 在 $\mathbb{R}$ 中找个反例.

1.5 证明引理 1.4 的推论.

1.6 证明 $p \in \overline{A}$ 当且仅当 $\rho(p, A) = 0$.

1.7 设 X 和 Y 为度量空间, 设 $f : X \to Y$ 具有性质: 只要 B 在 Y 中为开则 $f^{-1}[B]$ 在 X 中为开. 证明 f 为连续.

1.8 证明 $\rho(x, A) = \rho(x, \overline{A})$.

1.9 证明 $\rho(x, A)$ 是 x 的连续函数.(事实上它是利普希茨连续的.)

1.10 创建各种度量空间 S(选取 $\mathbb{R}^2$ 中子集), 它们各具下列性质:

(1) S 有 n 个点.

(2) S 为无限集且当 $x \neq y$ 时 $\rho(x, y) \geqslant 1$.

(3) S 有一个球 $B_1(a)$, 使得闭球 $\{x : \rho(x, a) \leqslant 1\}$ 与 $B_1(a)$ 的闭包不是一回事.

1.11 让明在一个赋范线性空间中一个闭球是其相应开球的闭包.

1.12 证明如果 $f : X \to Y$ 和 $g : Y \to Z$ 为连续 (其中 X, Y 和 Z 是度量空间), 则 $g \circ f$ 也如此.

1.13 设 X 和 Y 为度量空间. 定义在 $Z = X \times Y$ 上的乘积度量的概念. 定义 Z 上的 1 度量 ρ_1 和一致度量 ρ_∞(证明它们是度量). 这些定义类比于赋范线性空间的乘积空间上的 1 - 范数和一致范数, 并证明按你在前面的定义, 它们每一个都是乘积度量.

1.14 在 $Z = X \times Y$ 上对 2 - 度量做同样的事.

1.15 设 X 和 Y 为度量空间, 并设 V 为赋范线性空间. 设 $f : X \to \mathbb{R}$ 和 $g : Y \to V$ 为连续映射. 证明

$$\prec x, y \succ \mapsto f(x)f(y)$$

是从 $X \times Y$ 到 V 的连续映射.

*4.2 拓扑

如果 X 为一个任意集合, 而 $\mathfrak{T}$ 是 X 中满足定理 1.1 中从 (1) 到 (3) 性质的任一族子集合, 则称 $\mathfrak{T}$ 是 X 上的一个*拓扑*. 定理 1.1 断言说一个度量空间 X 的开子集形成了 X 上的一个拓扑. 接着发生的内部, 闭集, 和闭包的定义是纯粹拓扑的, 这个意思是说它们只依赖于拓扑 $\mathfrak{T}$, 定理 1.2 和等式 $(\overline{A})' = (A')^{\text{int}}$ 也是如此. 研究由于拓扑的存在而得到的结果被称为一般*拓扑学*.

另一方面，早先所给出的球和连续性的定义是*度量*定义，因此属于*度量*空间理论的部分. 于是，在度量空间中，我们不仅有拓扑，还有我们的连续性的 ε 定义和球，和闭包与内部的球面特征.

现在被告之虽然连续性和收敛性由度量的方式定义它们，也有纯拓扑的特征从而是拓扑思想，可能会使读者感到惊讶. 如果一个人记住在度量空间中一个开集不是别的而只是球的并，那么上面的事实是容易被弄明白的. 我们有:

f 在 p 连续当且仅当对每个包含 $f(p)$ 的开集 A 存在一个包含 p 的开集 B 使得 $f[B] \subset A$.

这个涉及一个单个点 p 周围行为的*局部*条件用邻域的概念提出来更为流畅. 一个集合 A 是点 p 的一个*邻域*是说如果 $p \in A^{\mathrm{int}}$. 于是我们有:

f 在点 p 连续当且仅当对 $f(p)$ 的每个邻域 N, $f^{-1}[N]$ 是 p 的一个邻域.

最后，有一个整体连续性的干净漂亮的拓扑特性描述. 假设 S_1 和 S_2 为拓扑空间. 于是 $f: S_1 \to S_2$(处处) 连续当且仅当只要 A 为开则 $f^{-1}[A]$ 为开. 又，f 为连续当且仅当只要 B 为闭则 $f^{-1}[B]$ 为闭. 由于引理 1.4，这些条件并不令人吃惊.

4.3 序列的收敛性

除了转向度量空间理论的更为广泛的观点之外，我们也要在我们的这套工具中添进序列收敛性的概念，读者大概还能从他从前所碰到的微积分中想起它来. 为什么度量空间理论较一般拓扑学更简单和更直观的主要理由之一在于，差不多所有度量式的论证都可以借用序列收敛表述；而在本章我们将持之以恒地使用它，较之于其他的选择更偏好了它，并以此部分地弥补先前对此工具的忽视.

定义 3.1 *我们说无穷序列* $\{x_n\}$ *收敛于点* a *是指如果对每个* ε *存在一个* N *使得*

$$n > N \ \Rightarrow\ \rho(x_n, a) < \varepsilon.$$

我们也说当 n 趋向无穷时 x_n 趋向 a，并称 a 为此序列的极限. 我们用符号写成当 $n \to \infty$, $x_n \to a$，或者 $\lim_{n\to\infty} x_n = a$. 从形式上看，这个定义实际上等同于我们早先的函数收敛性的定义，这时有着平行的定理，我们在一种情况下使用的论证一般几乎逐字逐句地在另一种情形成立. 第三章的引理 1.1 的证明可以稍稍变动一下给出了下面的结果.

引理 3.1 如果 $\{\xi_i\}$ 和 $\{\eta_i\}$ 为赋范线性空间 V 中的两个序列, 则

$$\xi_i \to \alpha \text{ 和 } \eta_i \to \beta \Rightarrow \xi_i + \eta_i \to \alpha + \beta.$$

主要的差别在于我们现在要选取 N 为 $\max\{N_1, N_2\}$ 以替代选取 δ 为 $\min\{\delta_1, \delta_2\}$. 相似地,

引理 3.2 如果在 V 中 $\xi_i \to \alpha$, 而在 $\mathbb{R}$ 中 $x_i \to a$, 则 $x_i\xi_i \to a\alpha$.

像从前那样, 此定义开始于三个量词, $(\forall\varepsilon)(\exists N)(\forall n)$. 用球的词和 "几乎所有的 n" 的概念重新将定义措词我们能够得到多少更符合语言习惯的形式. 我们说 $P(n)$ 对几乎所有的n 为直接的是除去有限个数外如果 $P(n)$ 对所有的 n 为真, 或者等价地, 如 $(\exists N)(\forall n^{>N})P(n)$. 于是我们看到

$\lim x_n = a$ 当且仅当在 a 的每个球包含了几乎所有的 x_n.

下面的序列式特征描述给出了大概是最直观的看待闭包和闭集概念的方式.

定理 3.1 点 x 在集合 A 的闭包 $\overline{A}$ 中当且仅当存在 A 中收敛于 x 的序列 $\{x_n\}$.

因此一个集合 A 为闭当且仅当在 A 中的每个收敛序列的极限在 A 中.

证明 如果 $\{x_n\} \subset A$ 并且 $x_n \to x$, 则在 x 的每个球包含了几乎每个 x_n, 因而特别地相交于 A. 因此由引理 1.3, $x \in \overline{A}$. 反之, 如果 $x \in \overline{A}$, 在 x 的每个球与 A 都相交, 从而我们可以在 $B_{1/n}(x) \cap A$ 中选任意点当作 x_n 以构造一个 A 中收敛于 x 的序列. 由于 A 为闭当且仅当 $A = \overline{A}$, 那么, 定理的第二个陈述可由第一个得出. □

也有一个连续性的序列特征描述, 它大大有助于以一种灵活的方式运用连续性概念. 设 X 和 Y 为度量空间. f 为 X 到 Y 的任意函数.

定理 3.2 函数 f 在 a 连续当且仅当, 对 X 中任意序列 $\{x_n\}$, 如果 $x_n \to a$ 则 $f(x_n) \to f(a)$.

证明 首先假设 f 在 a 连续, 并设 $\{x_n\}$ 为任意收敛于 a 的序列. 于是对给出的任意 ε, 存在一个 δ 使得

$$\rho(x, a) < \delta \Rightarrow \rho(f(x), f(a)) < \varepsilon,$$

这是 f 在 a 的连续性的意思; 对此 δ 存在一个 N, 由于 $x_n \to a$, 便有

$$n > N \;\Rightarrow\; \rho(x_n, a) < \delta.$$

结合这两个蕴涵式, 我们看到, 对给出的 ε 我们找到了 N, 使得 $n > N \Rightarrow \rho(f(x_n), f(a)) < \varepsilon$. 这即是 $f(x_n) \to f(a)$.

现假定 f 在 a 不连续. 在考虑这样一个否定词时, 重要之处在于把隐含的全称量词弄成明晰. 因此形式上, 我们假定了 $\sim (\forall\varepsilon)(\exists\delta)(\forall x)(\rho(x,a) < \delta \Rightarrow \rho(f(x), f(a)) < \varepsilon)$, 那就是说 $(\exists\varepsilon)(\forall\delta)(\exists x)(\rho(x,a) < \delta \;\&\; \rho(f(x), f(a)) \geqslant \varepsilon)$. 这种符号表示在读者还没有计算逻辑否定的实践时并不是必要的; 有经验的思考者不需要形式计算, 凭直觉就能找到正确的否定. 无论如何, 我们现在有一个固定的 ε, 并对每个形如 $1/n$ 的 δ 可以设 x_n 为一个相应的 x. 于是我们有 $\rho(x_n, a) < 1/n$ 和 $\rho(f(x_n), f(a)) \geqslant \varepsilon$ 对所有 n 成立。第一个不等式表明 $x_n \to a$, 而第二个表明 $f(x_n) \nrightarrow f(a)$. 因此如果 f 在 a不连续则序列条件就不被满足. □

上述论证的类型被极其频繁使用, 几乎等于在适当场合的自动证明程序. 譬如说, 我们要证明 $(\forall x)(\exists y)(\forall z)P(x,y,z)$. 由归谬法, 我们假定此不真, 故有 $(\exists x)(\forall y)(\exists z) \sim P(x,y,z)$. 因此, 不去使用所有的数 y 而代之以让 y 通过某收敛于 0 的序列, 如像 $\{1/n\}$, 且对每个这样的 y 选取一个对应的 z 或 z_n. 我们结束于 $\sim P(x, \frac{1}{n}, z_n)$ 对给出的 x 和所有的 n 成立, 从而以序列式的论证结束证明.

读者应记起在一个向量空间 V 上两个范数 p, q 等价当且仅当恒同映射 $\xi \mapsto \xi$ 为从 $\prec V, p \succ$ 到 $\prec V, q \succ$ 和 $\prec V, q \succ$ 到 $\prec V, p \succ$ 是连续的. 借助于上面的定理我们现在有:

定理 3.3 *范数 p 和 q 等价当且仅当它们产生出完全相同的收敛序列.*

我们早先曾论证说, 在两个赋范线性空间的积空间 $V \times W$ 上的范数应该等价于 $\| \prec \alpha, \xi \succ \|_1 = \|\alpha\| + \|\xi\|$. 现在关于这个和范数显然可以看出, 一个序列 $\prec \alpha_n, \xi_n \succ \in V \times W$ 收敛于 $\prec \alpha, \xi \succ$ 当且仅当在 V 中 $\alpha_n \to \alpha$ 且在 W 中 $\xi_n \to \xi$. 我们现在 (再次由定理 3.2) 看出了:

定理 3.4 *$V \times W$ 上的一个乘积范数是任何一个具有下面性质的范数: 在 $V \times W$ 中 $\prec \alpha_n, \xi_n \succ \to \prec \alpha, \xi \succ$ 当且仅当在 V 中 $\alpha_n \to \alpha$ 而在 W 中 $\xi_n \to \xi$.*

习题

3.1 证明一个度量空间中的收敛序列有一个惟一的极限. 就是说, 证明如果 $x_n \to a$ 和 $x_n \to b$, 则 $a = b$.

3.2 证明在度量空间中 $x_n \to x$ 当且仅当在 $\mathbb{R}$ 中 $\rho(x_n, x) \to 0$.

3.3 证明如果在 $\mathbb{R}$ 中 $x_n \to a$ 且对所有 n, $x_n \geqslant 0$ 则 $a \geqslant 0$.

3.4 证明如果在 $\mathbb{R}$ 中 $x_n \to 0$ 且对所有 n 有 $|y_n| \leqslant x_n$, 则 $y_n \to 0$.

3.5 给出引理 3.1 和 3.2 的详细的 (ε, N) 证明.

3.6 应用定理 3.2 证明, 如果 X 为度量空间, V 为赋范线性空间, 而 F 和 G 是从 X 到 V 的连续映射, 则 $F + G$ 连续. 叙述并证明对乘积 FG 的相似定理.

3.7 应用定理 3.2 证明在复合下连续性被保持.

3.8 证明在一度量空间中一个点的序列 (的值域) 一般来说不是个闭集. 证明它们也可以是个闭集.

3.9 在一个赋范线性空间中一个开球的闭包包含了相应闭球的这个事实实际上在引理 3.2 和定理 3.1 的基础上说来是微不足道的小事. 证明其确是如此.

3.10 直接证明, 如果在 $V = V_1 \times V_2$ 上使用极大范数 $\| \prec \alpha, \xi \succ \| = \max\{\|\alpha\|, \|\xi\|\}$, 则在 V 上

$$\prec \alpha_n, \xi_n \succ \to \prec \alpha, \beta \succ$$

成立当且仅当

$$\text{在 } V_1 \text{上 } \alpha_n \to \alpha \text{ 和在 } V_2 \text{ 上 } \xi_n \to \xi.$$

3.11 证明如果在 $\mathbb{R}^2$ 上 $\| \; \|$ 是个递增范数 (见第三章的定理 4.4 后面的评注), 则

$$\rho(\prec x_1, y_1 \succ, \prec x_2, y_2 \succ) = \| \prec \rho(x_1, x_2), \rho(y_1, y_2) \succ \|$$

是两个度量空间 X 和 Y 的积空间 $X \times Y$ 上的一个度量.

3.12 在上一个习题中证明在 $X \times Y$ 中 $\prec x_n, y_n \succ \to \prec x, y \succ$ 当且仅当在 X 中 $x_n \to x$ 而在 Y 中 $y_n \to y$. 这个性质或许是我们对一个*乘积度量*的最低要求.

3.13 定义一个乘积度量如上个习题, 用定理 3.2 证明

$$\prec f, g \succ : S \to X \times Y$$

连续当且仅当 $f : S \to X$ 和 $g : S \to Y$ 都连续.

3.14 设 X, Y 和 Z 为度量空间, 并设 $f : X \times Y \to Z$ 为一映射使得 $f(x, y)$ 分别对每个变量连续. 又假设对 x 的连续性*在 y 上是一致的*. 那就是说, 对给出的 ε 和 x_0, 存在一

个 δ 使得对每个 y 的值都有

$$\rho(x,x_0)<\delta \ \Rightarrow \ \rho(f(x,y),f(x,y))<\varepsilon.$$

证明 f 在 $X\times Y$ 上连续.

3.15 在闭单位正方形 $[0,1]\times[0,1]$ 上定义函数 f 为

$$f(0,0)=0, \text{如果} \prec x,y\succ\neq\prec 0,0\succ, \text{则 } f(x,y)=\frac{xy}{(x+y)^2}.$$

于是对每个固定的 y 值, f 作为 x 的函数是连续的, 反之也如此. 但是试证明 f 在原点是不连续的. 那就是说, 找出一个在平面上收敛于 $\prec 0,0\succ$ 的序列 $\prec x_n,y_n\succ$ 使得 $f(x_n,y_n)$ 不收敛于 0. 这个例子表明一个二元函数的连续性是一个比分别对每个变元连续更强的性质.

4.4 列紧性

读者应大体上熟悉子序列的概念. 一个序列 $\{x_n\}$ 的一个子序列是个新的序列 $\{y_n\}$, 它是由选择项 x_n 中无穷多个但不是全部, 并将它们按所选的指数编号. 因此如果 n_1 是第一个被选的 n, n_2 是其次的, 等等, 并且如果令 $y_m=x_{n_m}$, 我们便得到子序列

$$\{y_1,y_2,\cdots,y_m,\cdots\}=\{x_{n_1},x_{n_2},\cdots,x_{n_m},\cdots\}$$

或者

$$\{y_m\}_m=\{x_{n_m}\}_m.$$

严格地说, 对所选指标 n 集合的编号是个从 $\mathbb{Z}^+$ 到 $\mathbb{Z}^+$ 的序列 $m\mapsto n_m$, 它保持了顺序: $n_{m+1}>n_m$ 对所有 m 成立. 而子序列 $m\mapsto x_{n_m}$ 是序列 $n\mapsto x_n$ 与选择序列的复合.

为了避免下标的下标, 我们可使用 $n(m)$ 来代替 n_m. 在这两种情形的任一种中, 我们依习惯采取不甚认真的态度: 这同一符号 'n' 当我们写出 x_n 时它是一个整数值变量, 而当我们写出 $n(m)$ 或 n_m 时则当作选择函数. 这是标准的记号糊涂之一, 我们在初等微积分中已经忍受过了, 这是因为治好它比继续得病更坏. 我们能这样说: 设 f 是一个序列, 即一个从 $\mathbb{Z}^+$ 到 $\mathbb{R}$ 的一个函数. 于是 f 的一个子序列是一个复合 $f\circ g$, 其中 g 是从 $\mathbb{Z}^+$ 到 $\mathbb{Z}^+$ 的映射使得对所有的 m 有 $g(m+1)>g(m)$.

如果你已经掌握了子序列的思想, 你便应该能够看出任何一个 0 和 1 构成的无穷序列, 譬如 $\{0,1,0,0,1,0,0,0,1,\cdots\}$ 都可以作为子序列从 $\{0,1,0,1,0,\cdots,[1+(-1)^n]/2,\cdots\}$ 中得到.

如果 $x_n \to a$, 则应该清楚, 每个子序列也收敛于 a. 我们把此细节留给读者作为习题. 另一方面, 如果序列 $\{x_n\}$不收敛于a, 则有一个 ε, 使得对每个 N 有某个较大的 n, 在此 $\rho(x_n,a) \geqslant \varepsilon$. 现在我们可以对每个 N 选取一个 n, 并留意让 $n_{N+1} > n_N$, 因而选取了一个子序列, 它的所有的项离开 a 的距离至少为 ε. 因此 (我们刚刚选出的)这个序列没有收敛于 a 的子序列. 于是, 如果 $\{x_n\}$不收敛于 a, 则它有一个子序列, 这个子序列没有 (子) 子序列收敛于 a. 由此有

引理 4.1 如果序列 $\{x_n\}$ 和点 a 使得 $\{x_n\}$ 的每个子序列有它自身的收敛于 a 的子序列, 则 $x_n \to a$.

这是一个不合常理, 未必使人有深刻印象的引理, 但我们将用它去证明一个最重要的定理 (定理 4.2).

定义 4.1 度量空间的一个子集 A 为 (序)列紧的是说, 如果 A 中每个序列都有一个子序列收敛于 A 中一个点.

这里可以说, 我们无中生有地创造出了收敛性. 人们会期待一个紧集具有非常强的性质, 或许会怀疑并不存在很多的这种集合. 然而我们马上就会看到, $\mathbb{R}^n$ 中每个有界闭集都是紧的, 并且它是在有限维空间的理论中我们最经常使用的概念. 在无穷维空间中列紧性是一个极少有的现象, 但是当它确实存在时便非常重要, 正如我们将在第六章的施图姆 - 刘维尔的简短浏览中看到这一点.

我们从几个简单而重要的一般结果开始.

引理 4.2 如果 A 是度量空间 S 的一个列紧子集, 则 A 是闭和有界的.

证明 假设 $\{x_n\} \subset A$ 且 $x_n \to b$. 由 A 的紧性存在一个子序列 $\{x_{n(i)}\}_i$ 收敛于点 $a \in A$. 但一个收敛序列的子序列收敛于同一极限. 因此, $a = b$ 从而 $b \in A$. 故 A 为闭.

这里有界性意味着位于在一给定点 b 的某个球内. 如果 A 不是有界的, 那么对每个 n 存在一点 $x_n \in A$ 使得 $\rho(x_n,b) > n$. 由紧性, 子序列 $\{x_{n(i)}\}_i$ 收敛于一点 $a \in A$, 而且 $\rho(x_{n(i)},b) \to \rho(a,b)$. 这显然与 $\rho(x_{n(i)},b) > n(i) \geqslant i$ 矛盾. □

连续函数把紧集带到紧集. 下面结果的证明留给读者作习题.

定理 4.1 如果 f 连续, A 为它的定义域中一个列紧子集, 则 $f[A]$ 是列紧的.

一个非空紧集 $A \subset \mathbb{R}$ 包含了极大和极小元素. 这是因为 lub$A$ 是 A 中一个序列的极限, 因 A 为闭从而属于 A 本身. 结合这个事实和上面的定理, 我们得到下面熟悉的推论.

推论 4.1 如果 f 为连续实函数且 $\mathrm{dom}(f)$ 为非空和列紧, 则 f 有界并取得极大和极小值.

下面非常有用的结果与上面的定理有关.

定理 4.2 如果 f 连续并为双射, 而且 $\mathrm{dom}(f)$ 为列紧, 则 f^{-1} 连续.

证明 我们必须证明如果在 f 的值域中 $y_n \to y$, 并且如果 $x_n = f^{-1}(y_n)$, $x = f^{-1}(y)$, 则 $x_n \to x$. 只要证明每个子序列 $\{x_{n(i)}\}_i$ 自身有一个收敛于 x 的子序列就可以了 (引理 4.1). 但由于 $\mathrm{dom}(f)$ 为紧, 故存在子序列 $\{x_{n(i(j))}\}_j$ 收敛于某个 z, 而 f 的连续性蕴涵 $f(z) = \lim_{j\to\infty} f(x_{n(i(j))}) = \lim_{j\to\infty} y_{n(i(j))} = y$. 因此, $z = f^{-1}(y) = x$. 这正是我们所要证明的. 因此 f^{-1} 连续. □

我们现在着手解决证明 $\mathbb{R}^n$ 中有界闭集为紧的问题. 我们首先证明对 $\mathbb{R}$ 自身它成立, 然后对 $\mathbb{R}^n$ 给出一个归纳的论证.

称一个序列 $\{x_n\}$ 是*递增*的是说对所有 n 如果 $x_n \leqslant x_{n+1}$. 如果对所有 n 有 $x_n < x_{n+1}$ 则说它是*严格递增的*. *递降序列*和*严格递降序列*的概念是显然的. 一个或是递增或是递降的序列被称作单调的. 这些概念的意义在于下面的两个引理.

引理 4.3 $\mathbb{R}$ 中一个有界单调序列为收敛的.

证明 假设 $\{x_n\}$ 递增且上有界. 设 l 为其值域的上确界. 那就是说, $x_n \leqslant l$ 对所有 n 成立, 但对每个 $\varepsilon, l-\varepsilon$ 不是一个上界, 故对某个 N 有 $l-\varepsilon < x_N$. 于是

$$n > N \ \Rightarrow \ l-\varepsilon < x_N \leqslant x_n \leqslant l,$$

从而 $|x_n - l| < \varepsilon$. 即当 $n \to \infty$ 时 $x_n \to l$. □

引理 4.4 $\mathbb{R}$ 中任一序列都有一个单调的子序列.

证明 称 x_n 为一个峰项是说如果它大于或等于所有它后面的项. 如果有无穷多个峰项, 它们则显然形成了一个递降子序列. 另一方面, 如果只有有限个峰

项, 则有最后一个 x_n.(或者根本就没有), 那么每个后面的项都严格小于某个其他仍在更后面的项. 我们选取任意大于 n_0. 的任意 n_1, 而后再选 $n_2 > n_1$ 使 $x_{n_1} < x_{n_2}$ 等等. 因此这时我们选出了一个严格递增的子序列. 因此我们已经证明了 $\mathbb{R}$ 中任一序列 $\{x_n\}$ 或者有一个递降的子序列或者一个严格递增的子序列. □

将这两个引理合在一起, 我们有:

定理 4.3 *$\mathbb{R}$ 中每个有界序列有一个收敛的子序列. 现在我们可以用归纳法推广到 $\mathbb{R}^n$.*

定理 4.4 *$\mathbb{R}^n$ 中每个有界序列都有一个收敛的子序列 (使用任一个乘积范数, 譬如 $\|\ \|_1$).*

证明 前面一个定理是 $n = 1$ 的情形. 那么假设定理对 $n - 1$ 为真, 并设 $\{\boldsymbol{x}^m\}_m$ 为 $\mathbb{R}^n$ 中一个有界序列. 设想 $\mathbb{R}^n$ 为 $\mathbb{R}^{n-1} \times \mathbb{R}$, 我们有 $\boldsymbol{x}^m = \prec \boldsymbol{y}^m, z_m \succ$, 而 $\{\boldsymbol{y}^m\}_m$ 在 $\mathbb{R}^{n-1}$ 中有界, 这是因为如果 $\boldsymbol{x} = \prec \boldsymbol{y}, z \succ$, 则 $\|\boldsymbol{x}\|_1 = \|\boldsymbol{y}\|_1 + |z| \geqslant \|\boldsymbol{y}\|_1$. 因此由归纳假设, 有一个子序列 $\{\boldsymbol{y}^{n(i)}\}_i$ 收敛于 $\mathbb{R}^{n-1}$ 中某个 $\boldsymbol{y}$. 由于 $\{z_{n(i)}\}$ 在 $\mathbb{R}$ 中有界, 故它有一个子序列 $\{z_{n(i(p))}\}_p$ 收敛于 $\mathbb{R}$ 中某个 z. 当然, 相应的子子序列 $\{\boldsymbol{y}^{n(i(p))}\}_p$ 依然收敛于 $\mathbb{R}^{n-1}$ 中的 $\boldsymbol{y}$, 从而 $\{\boldsymbol{x}^{n(i(p))}\}_p$ 收敛于 $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$ 中的 $\boldsymbol{x} = \prec \boldsymbol{y}, z \succ$, 这是因为它的两个分量序列现在各自收敛于 $\boldsymbol{y}$ 和 z. 我们因此找到了一个收敛的 $\{\boldsymbol{x}^n\}$ 的子序列. □

定理 4.5 *如果 A 为 $\mathbb{R}^n$ 中的有界闭子集, 则 A 为列紧 (在任何一个范数下).*

证明 如果 $\{x_n\} \subset A$, 则由定理 4.4 存在一个子序列 $\{x_{n(i)}\}_i$ 收敛于 $\mathbb{R}^n$ 中某个 x, 并且因为 A 为闭集, x 在 A 中. 于是 A 为紧. □

我们现在来填补上一章中几个少数缺口中的一个.

定理 4.6 *$\mathbb{R}^n$ 中所有的范数都等价.*

证明 只要证明任一范数 $\|\ \|$ 等价于 $\|\ \|_1$ 就足够了. 令 $a = \max\{\|\boldsymbol{\delta}^i\|\}_1^n$, 我们有

$$\|\boldsymbol{x}\| = \left\|\sum_1^n x_i \boldsymbol{\delta}^i\right\| \leqslant \sum_1^n |x_i| \|\boldsymbol{\delta}^i\| \leqslant a\|\boldsymbol{x}\|_1,$$

故我们要证的不等式中的一个是平凡的. 我们也有 $|\ \|\boldsymbol{x}\| - \|\boldsymbol{y}\|\ | \leqslant \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant$

$a\|\boldsymbol{x}-\boldsymbol{y}\|_1$, 故 $\|\boldsymbol{x}\|$ 是 $\mathbb{R}^n$ 上相对于 $\|\ \|_1$ 的连续函数. 现在单位 1 - 范数球面 $S=\{\boldsymbol{x}:\|\boldsymbol{x}\|_1=1\}$ 是闭和有界的, 故而紧 (在 1 - 范数下). 连续函数 $\|\boldsymbol{x}\|$ 在这个紧集 S 上的限制有一个极小值 m, 而且 m 不为零, 这是因为 S 不包含零向量. 我们因而在 S 上有 $\|\boldsymbol{x}\|\geqslant m\|\boldsymbol{x}\|_1$, 故而按齐性, 在 $\mathbb{R}^n$ 上有 $\|\boldsymbol{x}\|\geqslant m\|\boldsymbol{x}\|_1$. 总之我们已经找到了正常数 a 和 m 使得 $m\|\ \|_1\leqslant\|\ \|\leqslant a\|\ \|_1$. □

与坐标同构复合, 我们看出在任何一个有限维向量空间上的所有范数都是等价的.

推论 4.2 如果 M 是赋范线性空间 V 中一个有限维子空间, 则 M 是 V 的闭子空间.

证明 假设 $\{\xi_n\}\subset M$ 且 $\xi_n\to\alpha\in V$. 我们必须证明 α 在 M 中. 现在 $\{\xi_n\}$ 是 M 中的一个有界子集, 而它在 M 中的闭包由定理是列紧的. 因此某个子序列收敛于 M 中一个点 β 但同样也收敛于 α, 故 $\alpha=\beta\in M$. □

习题

4.1 用归纳法证明, 如果 $f:\mathbb{Z}^+\to\mathbb{Z}^+$ 使得对所有 n 有 $f(n+1)>f(n)$, 则对所有 n 有 $f(n)\geqslant n$.

4.2 仔细地证明如果当 $n\to\infty$ 时 $x_n\to a$, 则当 $m\to\infty$ 时 $x_{n(m)}\to a$ 对任意子序列成立. 在这个证明中上一个习题是有用的.

4.3 证明如果 $\{x_n\}$ 是 $\mathbb{R}$ 中的一个递增序列(对所有 n, $x_{n+1}\geqslant x_n$), 并且如果 $\{x_n\}$ 中有一个收敛子序列, 则 $\{x_n\}$ 收敛.

4.4 对于陈述: 如果序列 $\{x_n\}$ 不收敛于 a, 则存在一个 ε 和一个子序列 $\{x_{n(m)}\}_m$ 使得对所有 m 有 $\rho(x_{n(m)},a)\geqslant\varepsilon$, 给出一个更仔细的论断.

4.5 找出 $\mathbb{R}$ 中一个没有收敛子序列的序列.

4.6 找出 $\mathbb{R}$ 中一个不收敛的序列使得收敛子序列的极限点集恰好由数 1 组成.

4.7 证明在 [0,1] 中有一个序列 $\{x_n\}$ 使得对任意 $y\in[0,1]$ 存在一个子序列 x_{n_m} 收敛于 y.

4.8 证明在一个度量空间 X 中一个序列 $\{x_n\}$ 的收敛子序列的极限的集合是 X 的闭子集.

4.9 证明定理 4.1.

4.10 证明两个列紧度量空间的笛卡儿积是列紧的 (其证明本质上同于正文中的.)

4.11 称一个度量空间是有界紧的是说, 如果每个闭的有界集是列紧的. 证明两个有界紧的度量空间的笛卡儿积仍是有界紧的 (使用譬如在乘积空间上的最大度量).

4.12 证明一个赋范线性空间中两个列紧子集的和 $A+B$ 仍是列紧的.

4.13 证明一个闭集和一个紧集的和 $A+B$ 为闭.

4.14 用 $\mathbb{R}$ 中的例子证明两个闭集的和不必是闭的.

4.15 设 $\{C_n\}$ 为列紧度量空间 S 中非空闭子集的逆降序列 (对所有 n, $C_{n+1}\subset C_n$). 证明 $\bigcap_1^\infty C_n$ 非空.

4.16 给出一个度量空间中非空闭子集的递降序列 $\{C_n\}$ 使得 $\bigcap_1^\infty C_n=\varnothing$ 的例子.

4.17 假设度量空间 S 具如后性质, 即 S 的每个非空闭子集的递降序列 $\{C_n\}$ 有非空交集. 证明 S 必是列紧的.

[提示: 给出任意序列 $\{x_i\}\subset S$, 并设 C_n 为 $\{x_i : i\geqslant n\}$ 的闭包.]

4.18 设 A 为一个 nls$V$ 的列紧子集, 并设从 A 中点到原点画出的所有线段得到的集合为 B(就是说,

$$B=\{t\alpha:\alpha\in A \quad 和 \quad t\in[0,1]\}).$$

证明 B 为紧.

4.19 把一个紧性的论断用在引理 1.5, 证明如果 N 是一个有限维向量空间 V 的真闭子空间, 则存在 V 中 α 使得 $\|\alpha\|=\rho(\alpha,N)=1$.

4.5 紧性和一致性

'一致' 这个词在数学中常常用作一个定性的形容词. 粗略地说, 它关系到一个 "点" 性质 $P(y)$, 它对一个区域 A 中每个点 y 可能成立也可能不成立, 而它的定义涉及一个存在量词. 对 $P(y)$ 的一个典型形式是 $(\forall c)(\exists d)Q(y,c,d)$. 因此, 如果 $P(y)$ 是 'f 在 y 为连续' 则 $P(y)$ 有形式 $(\forall\varepsilon)(\exists\delta)Q(y,\varepsilon,\delta)$. 此性质在 A 上成立是说对 A 中所有的 y 成立, 就是说, 如果

$$(\forall y^{\in A})[(\forall c)(\exists d)Q(y,c,d)].$$

一般来说, 这里的 d 同时依赖于 y 和 c; 如果 y 或 c 两者之一被改变, 则相应的 d 会不得不变化. 因此在连续性定义中的 δ 同时依赖于 ε 和那个在其中断言有连续性的点 y. 称这个性质在 A 上一致地或对 y 一致地成立是说如果可以找到一个与 y 无关的值 d(但仍与 c 相关). 因此这性质对 y 一致地成立是说, 如果

$$(\forall c)(\exists d)(\forall y^{\in A})Q(y,c,d);$$

此性质的一致性是以量词 $(\forall y^{\in A})$ 和 $(\exists d)$ 的反顺序来表达的. 因此 f 在 A 上是一致连续的是说如果

$$(\forall \varepsilon)(\exists \delta)(\forall y, z^{\in A})[\rho(y,z) < \delta \ \Rightarrow\ \rho(f(y), f(z)) < \varepsilon] .$$

现在 δ 与我们断言在那里有连续性的点无关, 但当然仍旧依赖于 ε.

我们在上一章第 14 节中看到当连续性的点条件一致地成立时它是多么更加强大有力. 在本节后面部分我们将讨论一些其他的一致性概念, 并且我们将看到, 如果点性质成立的区域是列紧的, 则此点性质常常蕴涵了一致性质.

我们在上面已考查过的形式的叙述形式清楚地表明了一致性和非一致性之间的不同特征. 但在写出一个论证时, 我们一般会按照我们更为符合语言习惯的用法, 省去了内部的全称量词. 例如, 一个函数序列 $\{f_n\} \subset W^A$ 逐点收敛于 $f: A \to W$ 是说, 如果它在 A 中每点 p 收敛于 f, 那就是, 如果对 A 中每点 p 和每个 ε, 存在一个 N 使得对每个 A 中的 p,

$$n > N \ \Rightarrow\ \rho(f_n(p), f(p)) \leqslant \varepsilon.$$

当 $\rho(\xi, \eta) = \|\xi - \eta\|$, 说 $\rho(f_n(p), f(p)) \leqslant \varepsilon$ 对所有 p 成立与说 $\|f_n - f\|_\infty \leqslant \varepsilon$ 是一样的. 因此一致地 $f_n \to f$ 当且仅当 $\|f_n - f\|_\infty \to 0$; 这就是为什么称 $\|f\|_\infty$ 为一致范数的理由.

逐点的收敛性并不意味一致收敛性. 因此, $f_n(x) = x^n$ 在 $A = (0,1)$ 上逐点收敛于零函数但不是一致收敛.

在 A 上的连续性也不表明是一致连续的. 函数 $f(x) = 1/x$ 在 $A = (0,1)$ 连续但不是一致连续. 然而紧性改变了后面的这种情况.

定理 5.1 *如果 f 在 A 上连续, 并且 A 为紧集, 则 f 在 A 上一致连续.*

证明 这是我们的 "自动的" 否定证明之一. 一致连续性 (UC) 是性质

$$(\forall \varepsilon^{>0})(\exists \delta^{>0})(\forall x, y^{\in A})[\rho(x,y) < \delta \ \Rightarrow\ \rho(f(x), f(y)) < \varepsilon].$$

因此 , $\sim$UC$\Leftrightarrow (\exists \varepsilon)(\forall \delta)(\exists x, y)[\rho(x,y) < \delta$ 和 $\rho(f(x), f(y)) \geqslant \varepsilon]$. 取 $\delta = \frac{1}{n}$, 从而有对应的 x_n, y_n. 因此对所有 n 有 $\rho(x_n, y_n) < 1/n$ 和 $\rho(f(x_n), f(y_n)) \geqslant \varepsilon$, 其中 ε 为一个固定的函数. 现在 $\{x_n\}$ 由 A 的紧性有一个收敛的子序列, 譬如 $x_{n(i)} \to x$. 因为 $\rho(y_{n(i)}, x_{n(i)}) < \frac{1}{i}$, 我们也有 $y_{n(i)} \to x$. 由 f 在 x 的连续性, 有

$$\rho(f(x_{n(i)}), f(y_{n(i)})) \leqslant \rho(f(x_{n(i)}), f(x)) + \rho(f(x), f(y_{n(i)})) \to 0,$$

它与 $\rho(f(x_{n(i)}), f(y_{n(i)})) \geqslant \varepsilon$ 矛盾. 由否定完成了证明. □

但是 A 的紧性不是自动地把 A 上函数序列的逐点收敛性转换为一致收敛性. 由图 4.1 展示的图像定义的 "分段线性" 函数 $f_n:[0,1]\to[0,1]$ 在紧区域 $[0,1]$ 上逐点收敛于零, 但此收敛性不是一致的 (参看习题 5.4).

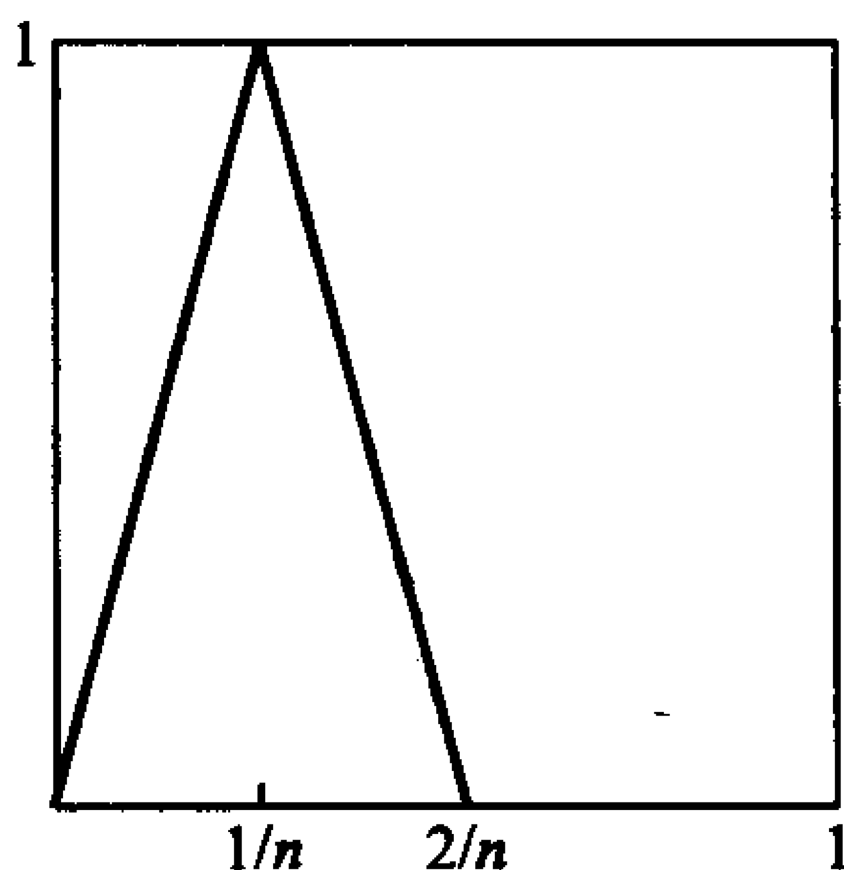


图 4.1

我们在前面曾指出一对不交的闭集间的距离可能为零. 但是如果其中一个闭集是紧的, 则此距离必为正.

定理 5.2 如果 A 和 C 为不变非空闭集, 且其中之一为紧, 则 $\rho(A,C)>0$.

证明 用自动的归谬法证明, 并将其留给读者. □

这个结果又是一个一致性条件. 说一个集合 A 与一个闭集 C 不交就是说 $(\forall x^{\in A})(\exists r^{>0})\ (B_r(x)\cap c=\varnothing)$. 说 $\rho(A,C)>0$ 就是说 $(\exists r^{>0})(\forall x^{\in A})\ \cdots$.

作为列紧性的最后一个结果, 我们将建立一个非常有威力的性质, 在一般拓扑学中将此性质作为紧性的定义. 但是我们首先需要做一些预备性的工作. 如果 A 是度量空间 S 中的一个子集, A 的 *r*邻域$B_r[A]$ 可以简单地说成是 A 中所有点的半径为 r 的球的并:

$$B_r[A]=\cup\{B_r(a):a\in A\}=\{x:(\exists a^{\in A})(\rho(x,a)<r)\}.$$

一个子集 $A\subset S$ 是 r -稠于 S 的是说如果 $S\subset B_r[A]$, 即如果 S 的每个点到 A 中某个点比 r 更近.

度量空间 S 的一个子集 A 是稠于 S 的是说如果 $\overline{A}=S$. 这与说对于每个 S 中点 p, 存在 A 中任意靠近 p 的点是一样的. 所有有理数的集合 $\mathbb{Q}$ 是实数系 $\mathbb{R}$ 的稠子集, 这是因为任意一个无理实数 x 可以被有理数任意地靠近. 由于我们用十进制记号做算术, 故习惯上用使用小数近似, 于是如果 $0<x<1$ 且 x 的小

数展开为 $x = \sum_1^\infty a_n/10^n$, 其中每个 a_n 是 $0 \leqslant a_n < 10$ 的整数, 则 $\sum_1^N a_n/10^n$ 是个与 x 相差小于 10^{-N} 的有理数. 注意, A 是 B 的稠子集当且仅当 A 对所有正数 r 在 B 中为 r - 稠的.

称一个集合 B 是完全有界的是说如果对每个正数 r 存在一个有限集它 r - 稠于 B. 因此对每个正 r, 集合 B 可以被有限个半径为 r 的球覆盖. 例如, $n-1$ 个数 $\{i/n\}_1^{n-1}$ 是 $(1/n)$ - 稠于开区间 (0,1), 这对所有 n 都对, 故 (0,1) 是完全有界的.

像下面引理表明的那样, 完全有界性是比有界性要强很多的性质.

引理 5.1 如果赋范线性空间 V 是无限维的, 则它的闭单位球 $B_1 = \{\xi : \|\xi\| \leqslant 1\}$ 不能被有限个半径为 $\frac{1}{3}$ 的球所覆盖.

证明 因为 V 不是有限维的, 所以我们能够选取一个序列 $\{\alpha_n\}$ 使得对每个 n α_{n+1} 不在 $\{\alpha_1, \cdots, \alpha_n\}$ 线性张成的 M_n 中, 由于 M_n 闭于 V(定理 4.6 的推论), 我们可以应用引理 1.5 找出 M_n 中一个向量 ξ_n 使得 $\|\xi_n\| = 1$ 而 $\rho(\xi_n, M_{n-1}) > 2/3$ 对所有 $n > 1$ 成立. 取 $\xi_1 = \alpha_1/\|\alpha_1\|$, 我们便有了一个序列 $\{\xi_n\} \subset B_1$ 使得当 $m \neq n$ 时

$$\|\xi_m - \xi_n\| > \frac{2}{3}.$$

没有一个半径为 $\frac{1}{3}$ 的球能够包含多于一个的 ξ_i, 引理得证. □

举个具体的例子. 设 V 为 $\mathfrak{C}([0,1])$, f_n 为图 4.2 所描画的 "峰" 函数, 其中在底部的三个点是 $1/(2n+2), 1/(2n+1)$, 和 $1/2n$. 于是 f_{n+1} 与 f_n"不相交"(即 $f_{n+1}f_n = 0$), 并且我们对所有 n 有 $\|f_n\|_\infty = 1$, 以及 $\|f_n - f_m\|_\infty = 1$. 当 $n \neq m$. 因此没有半径为 1/2 的球可以包含多于一个的这种函数 f_n, 故在 V 中的单位闭

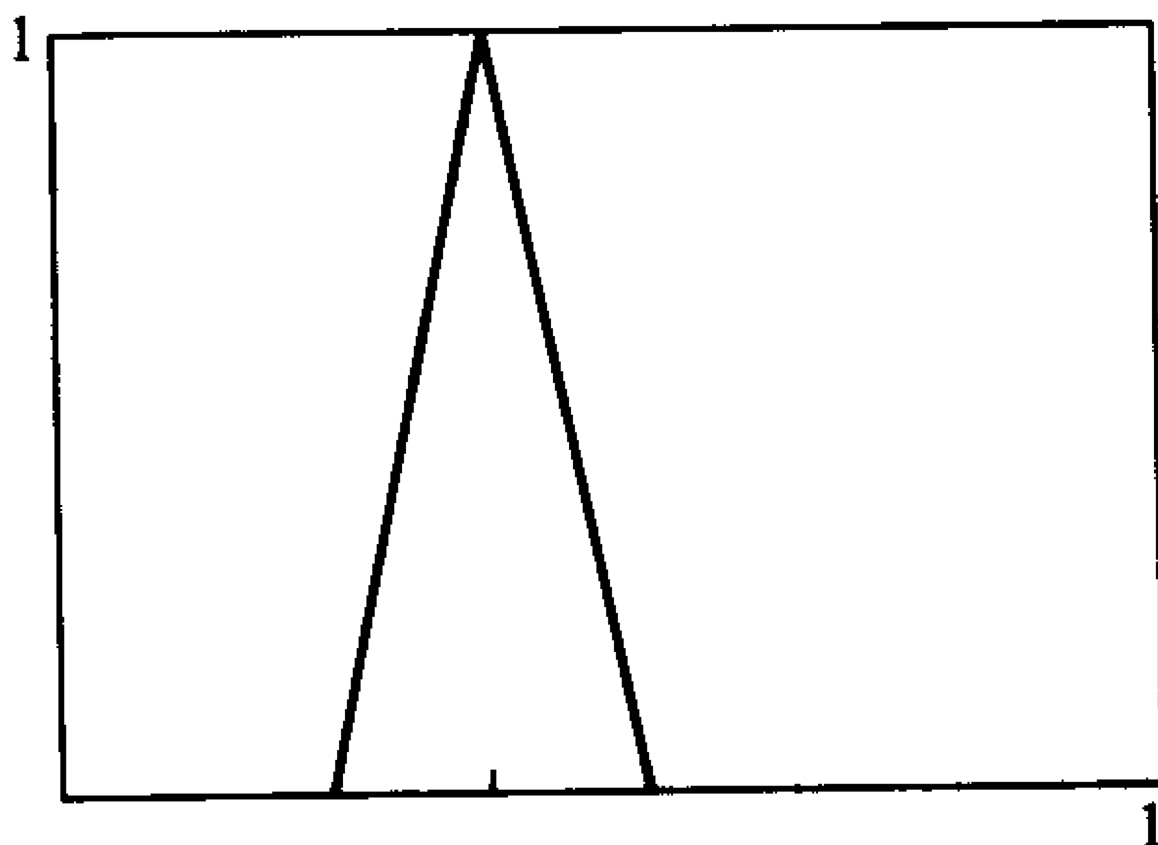


图 4.2

球不能被有限个半径为 1/2 的球覆盖.

引理 5.2 每个列紧集 A 是完全有界的.

证明 如果 A 不是完全有界的, 则存在一个 r 使得没有一个有限子集 F 在 A 中为 r - 稠的. 我们于是可以归纳定义一个序列 $\{p_n\}$ 使 p_1 为 A 中任意一个点, p_2 为一个不在 $B_r(p_1)$ 中的任一点, p_n 为不在 $B_r[\bigcup_1^{n-1} p_i] = \bigcup_1^{n-1} B_r(p_i)$ 中的任一点. 于是 $\{p_n\}$ 是 A 中一个序列使得 $\rho(p_i, p_j) \geqslant r$, $i \neq j$. 但这样的序列没有收敛的子序列. 因此, 如果 A 不是完全有界的则 A 便不是列紧的, 这证明了引理. □

推论 5.1 一个赋范线性空间是有限维的当且仅当它的闭单位球是列紧的.

证明 推论的一个方向的证明由定理 4.4 得到, 而另一个方向由上面两个引理得到. □

引理 5.3 假设 A 为列紧而 $\{E_i : i \in I\}$ 为 A 的一个开覆盖 (即 $\{E_i\}$ 是一族开集且 $A \subset \cup_i E_i$). 于是存在一个 $r > 0$ 使得对每个 A 中点 p, 球 $B_r(p)$ 被包含在某个 E_j 中.

证明 否则, 对每个 r 存在 A 中一个点 p 使得 $B_r(p)$ 不是任何一个 E_j 的子集. 取 $r = 1/n$, 得到相应的序列 $\{p_n\}$. 因此 $B_{1/n}(p_n)$ 不是任何一个 E_j 的子集. 由于 A 为列紧, 故 $\{p_n\}$ 有个收敛的子序列, 当 $m \to \infty$ 时 $p_{n(m)} \to p$. 由于 $\{E_i\}$ 覆盖了 A, 某个 E_j 包含 p, 于是对某个 $\varepsilon > 0$ 有 $\beta_\varepsilon(p) \subset E_j$(这是由于 E_j 为开集). 取 m 足够大使得 $1/m < \varepsilon/2$ 且 $\rho(p_{n(m)}, p) < \varepsilon/2$, 我们有

$$B_{1/n(m)}(p_{n(m)}) \subset B_\varepsilon(p) \subset E_j,$$

与 $B_{1/n}(p_n)$ 不是任一个E_j 的子集这个事实矛盾. 引理因此而得证. □

定理 5.3 如果 $\mathfrak{F}$ 是列紧集 A 的一个开覆盖, 则存在 $\mathfrak{F}$ 的一个有限子族覆盖了 A.

证明 由刚刚得到的引理知道, 存在一个 $r > 0$ 使得对 A 中每个 p, 球 $B_r(p)$ 完全在 $\mathfrak{F}$ 的某个集合内, 再由第一个引理, 在 A 中存在 $p_1, \cdots, p_n$ 使得 $A \subset \cup_1^n B_r(p_i)$. 取 $\mathfrak{F}$ 中相应的 E_i, 使 $B_r(p_i) \subset E_i, i = 1, \cdots, n$, 显然我们有 $A \subset \cup_1^n E_i$. □

在一般拓扑学中，一个集合 A 使得 A 的每个开覆盖都包含了一个有限覆盖被称作紧集合或说具有海涅 - 波雷尔性质. 上面定理说，在一个度量空间中每个列紧集是紧的. 我们在下面将看到相反的蕴涵关系也成立，故这两个概念事实上在度量空间中是等价的.

定理 5.4 *如果 A 是紧度量空间，则 A 是列紧的.*

证明 设 $\{x_n\}$ 为 A 中任一序列，设 $\mathfrak{F}$ 是那些开球 B 的一个汇集，其中每个 B 只包含有限个 x_i. 倘若 $\mathfrak{F}$ 覆盖了 A，则由紧性，A 被 $\mathfrak{F}$ 中有限多个球的并覆盖，这显然意味着整个 A 只包含了有限多个 x_i，这与 $\{x_i\}$ 是个无穷序列相矛盾. 因此 $\mathfrak{F}$ 没有覆盖 A，故存在 A 中一个点 x，使得在 x 的每个球包含了无穷多个这些 x_i. 更准确地说，在 x 的每个球包含了对无穷多个指标 i 的 x_i. 现在可以放心地留给读者去看出 $\{x_n\}$ 中有个子序列收敛于 x. □

习题

5.1 证明 $f_n(x) = x^n$ 在 (0,1) 上不是一致收敛的.

5.2 证明 $f(x) = 1/x$ 在 (0,1) 上不是一致连续的.

5.3 定义一个函数 $K: X \times Y \to Y$ 在第一个变量上对第二个变量是一致利普希茨的概念.

5.4 设 S 为一个列紧度量空间，并设 $\{f_n\}$ 是 S 上连续实函数的一个序列，它逐点地递降到 0(即 $\{f_n(p)\}$ 是 $\mathbb{R}$ 中的递降序列且当 $n \to \infty$ 时 $f_n(p) \to 0$). 证明此收敛性是一致的.(用习题 4.15 试一试.)

5.5 借助于本节定理 5.1 运用足够而又较弱的假设条件重新叙述第三章的定理 15.1 和 15.2 的推论.

5.6 证明定理 5.2.

5.7 证明如果 A 是赋范线性空间集合 X 的一个 r - 稠子集，并且如果 B 是集合 $Y \subset V$ 中一个 s - 稠子集，则 $A + B$ 是 $X + Y$ 中的 $(r + s)$ - 稠子集. 最后得到 V 的两个完全有界子集的和是完全有界的结论.

5.8 假设 n 个点 $\{p_i\}_1^n$ 在一个度量空间 X 中为 r - 稠. 设 A 为 X 的任一子集. 证明 A 中有一个最多为 n 个点的子集，它 $2r$ - 稠于 A. 最后得到一个完全有界度量空间的任何子集本身也是完全有界的结论.

5.9 证明两个完全有界度量空间的笛卡儿积是完全有界的.

5.10 证明如果一个度量空间 X 有一个完全有界的稠子集 A，则 X 为完全有界.

5.11 证明如果两个从度量空间 X 到度量空间 Y 的连续映射 f 和 g 在 X 的一个稠子集上相等, 则它们处处相等.

5.12 以涉及球的存在性的显式定量形式写出集合 $\{A_i\}$ 的内点集覆盖度量空间 A 的陈述. 然后证明引理 5.3 的结论是另一个一致性断言.

5.13 在定理 5.3 的基础上重新证明在一个紧的定义域上的一个连续函数是有界的这个定理.

5.14 由定理 5.3 重新证明在一个紧的定义域上的连续函数为一致连续这个定理.

4.6 等度连续性

我们将要做的是在无穷维背景下对列紧性的应用, 它涉及到*等度连续*的函数族的概念。如果 A 和 B 是度量空间, 则称一个子集 $\mathfrak{F} \subset B^A$ 在 A 中 p_0 为*等度连续*的是说, 如果 $\mathfrak{F}$ 的所有函数在 p_0 连续而且如果对给出的 ε, 存在一个 δ 对它们全部都适用, 就是说, 使得对 $\mathfrak{F}$ 中每个 f 都有

$$\rho(p, p_0) < \delta \;\Rightarrow\; \rho(f(p), f(p_0)) < \varepsilon.$$

族 $\mathfrak{F}$ 是*一致等度连续的*是说如果 δ 也与 p_0 无关, 从而只依赖于 ε. 我们的量词串因此是 $(\forall\varepsilon)(\exists\delta)(\forall p, q^{\in A})(\forall f^{\in \mathfrak{F}})$.

例如, 给出 $m > 0$, 设 $\mathfrak{F}$ 为从 (0,1) 到 (0,1) 的函数汇集, 使得 f' 存在且在 (0,1) 上 $|f'| \leqslant m$. 于是由普通的中值定理知 $|f(x) - f(y)| \leqslant m|x - y|$. 因此对给出的任意 ε, 我们可以取 $\delta = \varepsilon/m$ 并对所有 $x, y \in (0,1)$ 和所有 $f \in \mathfrak{F}$ 有

$$|x - y| < \delta \;\Rightarrow\; |f(x) - f(y)| < \varepsilon.$$

这个汇集 $\mathfrak{F}$ 从而为一致等度连续.

定理 6.1 *如果 A 和 B 为完全有界的度量空间, 并且如果 $\mathfrak{F}$ 是 B^A 中一个一致等度连续的子族, 则 $\mathfrak{F}$ 在一致度量下为完全有界.*

证明 给出 $\varepsilon > 0$, 选取 δ 使得对所有 $f \in \mathfrak{F}$ 和所有 $p_1, p_2 \in A$ 有 $\rho(p_1, p_2) < \delta \Rightarrow \rho(f(p_1), f(p_2)) < \varepsilon/4$. 设 D 为 A 中一个有限子集, 它 δ - 稠于 A, 并设 E 为 B 中一个有限子集, 它 $(\varepsilon/4)$ - 稠于 B. 设 G 为 D 到 E 中的所有函数的集合 E^D. G 自然是有限的; 事实上, $\#G = n^m$, 其中 $m = \#D, n = \#E$. 最后, 对每个 $g \in G$ 让 $\mathfrak{F}_g$ 为所有函数 $f \in \mathfrak{F}$ 使得对每个 $p \in D$ 有

$$\rho(f(p), g(p)) < \varepsilon/4$$

的集合. 我们断言这些 $\mathfrak{F}_g$ 覆盖了 $\mathfrak{F}$, 而且每个 $\mathfrak{F}_g$ 的直径最多为 ε. 由此我们将得到一个 $\mathfrak{F}$ 的有限 ε - 稠子集, 这只须从每个非空的 $\mathfrak{F}_g$ 中选取一个函数就可以做到, 从而使定理得证.

要证明每个 $f \in \mathfrak{F}$ 在某个 $\mathfrak{F}_g$ 中, 我们简单地构造一个适当的 g. 对每个 D 中的 p 存在 E 中的一个 q, 它与 $f(p)$ 的距离小于 $\varepsilon/4$. 如果我们对 D 中每个 p 选取一个在 E 中这样的 q, 则有了一个 G 中的函数 g 使得 $f \in \mathfrak{F}_g$.

我们必须做的最后一件事是证明如果 $f, h \in \mathfrak{F}_g$, 则 $\rho(f,g) \leqslant \varepsilon$. 由于 $\rho(h,g) < \varepsilon/4$, $\rho(f,g) < \varepsilon/4$ 在 D 上都成立, 则对每个 $p \in D$ 得到

$$\rho(f(p), h(p)) < \varepsilon/2.$$

于是对任意 $p' \in A$ 我们只须选取 $p \in D$ 使得 $\rho(p', p) < \delta$, 从而我们有

$$\begin{aligned}\rho(f(p'), h(p')) &\leqslant \rho(f(p'), f(p)) + \rho(f(p), h(p)) + \rho(h(p), h(p')) \\ &\leqslant \varepsilon/4 + \varepsilon/2 + \varepsilon/4 = \varepsilon.\end{aligned}$$ □

上面的证明是个数学论证的好例子, 它十分初等但艰难. 在提及数学推理时,'高级' 和 '困难' 这两个词绝不等价.

4.7 完备性

如果当 $n \to \infty$ 时 $x_n \to a$, 则这些项 x_n 当 n 变大时显然相互变得靠近. 另一方面, 如果 $\{x_n\}$ 是当 $n \to \infty$ 时它们的项可以相互靠得任意近的一个序列, 则 $\{x_n\}$ 显然应该收敛于一个极限, 但也可能不是; 这个期待的极限点可以不在此空间. 如果一个度量空间 S 使得每个应该收敛的每个序列实际上都收敛, 我们则说 S 是*完备的*. 我们现在使此概念变得准确清晰.

定义 7.1 *称* $\{x_n\}$ *是个柯西序列是说如果对每个* ε *存在一个* N *使得*

$$m > N \text{ 和 } n > N \ \Rightarrow\ \rho(x_m, x_n) < \varepsilon.$$

引理 7.1 *如果* $\{x_n\}$ *收敛, 则* $\{x_n\}$ *是柯西序列.*

证明 对给出的 ε, 我们选取 N 使得 $n > N \Rightarrow \rho(x_n, a) < \varepsilon/2$, 其中 a 为此序列的极限. 那么, 如果 m 和 n 同时都大于 n, 我们便有

$$\rho(x_m, x_n) \leqslant \rho(x_m, a) + \rho(a, x_n) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$ □

引理 7.2 如果 $\{x_n\}$ 是柯西的, 并且如果有一个子序列收敛, 则 $\{x_n\}$ 自身也收敛.

证明 假定当 $i \to \infty$ 时 $x_{n(i)} \to a$. 给出 ε, 我们取 N 使得 $m,n > N \Rightarrow \rho(x_n, x_m) < \varepsilon$. 因为当 $i \to \infty$ 时 $x_{n(i)} \to a$, 我们可以选一个 i 使得 $n(i) > N$ 及 $\rho(x_{n(i)}, a) < \varepsilon$. 因此如果 $m > N$, 我们则有

$$\rho(x_m, a) \leqslant \rho(x_m, x_{n(i)}) + \rho(x_{n(i)}, a) < 2\varepsilon,$$

故 $x_m \to a$. □

自然, 实际上如果 $m, n > N \Rightarrow \rho(x_m, x_n) < \varepsilon$, 并且如果 $x_n \to a$, 则对任意 $m > N$, $\rho(x_m, a) \leqslant \varepsilon$. 为什么?

引理 7.3 如果 A 和 B 为度量空间, 并且如果 T 是从 A 到 B 的利普希茨映射, 则 T 将 A 中的柯西序列带到 B 中的柯西序列. 特别, 如果 A 和 B 为赋范线性空间并且 T 是 $\mathrm{Hom}(A, B)$ 中元素时, 这是正确的.

证明 设 $\{x_n\}$ 为 A 中一个柯西序列, 并令 $y_n = T(x_n)$. 给出 ε, 选取 N 使得 $m, n > N \Rightarrow \rho(x_m, x_n) < \varepsilon/C$, 其中 C 是 T 的一个利普希茨常数. 于是

$$m, n > N \Rightarrow \rho(y_m, y_n) = \rho(T(x_m), T(x_n)) \leqslant C\rho(x_m, x_n) < C\varepsilon/C = \varepsilon. \quad \square$$

这个引理有如下的实质性推广.

定理 7.1 如果 A 和 B 为度量空间, $\{x_n\}$ 为 A 中的柯西序列, 而 $F: A \to B$ 为一致连续, 则 $\{F(x_n)\}$ 在 B 中是柯西的.

证明 证明作为习题留下. □

同学们应努力获取对这些引理真实性的良好直觉, 依照此感觉这些技术性证明变得多多少少有些显而易见.

定义 7.2 一个度量空间 A 是完备的是说如果 A 中的每个柯西序列都收敛于 A 中的一个极限. 称一个完备的赋范线性空间为一个巴拿赫空间.

我们现在要列举一些重要的巴拿赫空间的例子. 在每种情形给出一个证明是必要的, 故将这些例子写成一些定理.

定理 7.2 $\mathbb{R}$ 是完备的.

证明 设 $\{x_n\}$ 在 $\mathbb{R}$ 中为柯西的. 于是 $\{x_n\}$ 有界 (为什么?), 故由定理 4.3, 它有一个收敛的子序列. 引理 7.2 则意味着 $\{x_n\}$ 是收敛的. □

定理 7.3 如果 A 是个完备的度量空间, 并且如果 f 是从 A 到度量空间 B 的连续双射映射, 使得 f^{-1} 为利普希茨连续, 则 B 为完备. 特别, 如果 V 是个巴拿赫空间, 并且 T 在 $\mathrm{Hom}(V,W)$ 可逆, 则 W 为巴拿赫空间.

证明 假设 $\{y_n\}$ 是 B 中的一个柯西序列, 并令对所有的 i, $x_i = f^{-1}(y_i)$. 于是由引理 7.3, $\{x_i\}$ 在 A 中为柯西的, 故收敛于 A 中的某个 x, 这是因为 A 为完备的缘故. 但是因为 f 连续, 所以 $y_n = f(x_n) \to f(x)$. 因此 B 中每个柯西序列都收敛, 从而 B 为完备. □

巴拿赫空间的断言是个特殊情形, 因为 T 的可逆性意味着 T^{-1} 存在且在 $\mathrm{Hom}(W,V)$ 中, 从而是个利普希茨映射.

推论 7.1 如果 p 和 q 在 V 上为等价范数, 而且 $\prec V,p \succ$ 为完备, 则 $\prec V,q \succ$ 也如此.

定理 7.4 如果 V_1 和 V_2 为巴拿赫空间, 则 $V_1 \times V_2$ 也是.

证明 如果 $\{\prec \xi_n, \eta_n \succ\}$ 是柯西的, 则 $\{\xi_n\}$ 和 $\{\eta_n\}$ 中的每一个也是柯西的 (因为投射π_i 有界, 故由引理 7.3 可知). 于是 $\xi_n \to \alpha$ 和 $\eta_n \to \beta$, 其中 $\alpha \in V_1$ 和 $\beta \in V_2$. 因此在 $V_1 \times V_2$ 中 $\prec \xi_n, \eta_n \succ \to \prec \alpha, \beta \succ$.(见定理 3.4.) □

推论 7.2 如果 $\{V_i\}_1^n$ 为巴拿赫空间, 则 $\prod_{i=1}^n V_i$ 也是.

推论 7.3 每个有限维向量空间都是巴拿赫空间. (在任意范数下.)

证明 由定理 7.2 和上面的推论 1 知 $\mathbb{R}^n$ 为完备 (譬如说在 1 范数下). 然后我们在 V 上选择一组基并赋以 1 范数, 再应用定理 7.3 的推论把它过渡到其他任何一个范数上. □

定理 7.5 设 W 为巴拿赫空间, A 为任意集, 并设 $\mathfrak{B}(A,W)$ 为从 A 到 W 的所有有界映射构成的向量空间, 其范数为一致范数 $\|f\|_\infty = \mathrm{lub}\{\|f(a)\| : a \in A\}$. 于是 $\mathfrak{B}(A,W)$ 为巴拿赫空间.

证明 设 $\{f_n\}$ 为柯西序列并取任意 $a \in A$. 因为 $\|f_n(a) - f_m(a)\| \leqslant \|f_n - f_m\|_\infty$, 故 $\{f_n(a)\}$ 在 W 中是柯西的, 从而收敛. 定义 $g: A \to W$ 为 $g(a) = \lim f_n(a)$, 对 A 中每个 a 成立. 我们必须证明 g 为有界且 $f_n \to g$.

给出 ε, 我们选取 N 使得 $m, n > N \Rightarrow \|f_m - f_n\|_\infty < \varepsilon$.

$$\|f_m(a) - g(a)\| = \lim_{n \to \infty} \|f_m(a) - f_n(a)\| \leqslant \varepsilon.$$

于是, 如果 $m > N$, 则 $\|f_m(a) - g(a)\| \leqslant \varepsilon$ 对所有 $a \in A$ 成立, 从而 $\|f_m - g\|_\infty \leqslant \varepsilon$. 这表明 f_m 和 $f_m - g$ 都属于 $\mathfrak{B}(A, W)$, 所以

$$g = f_m - (f_m - g) \in \mathfrak{B}(A, W),$$

且在一致范数下 $f_m \to g$. □

定理 7.6 如果 V 是赋范线性空间, W 是个巴拿赫空间, 则 $\mathrm{Hom}(V, W)$ 是个巴拿赫空间.

其证明方法等同于前一个定理的, 故将其留作习题. 这里的有界性的意思不一样, 但以本质上一样的方式加以使用. 需要建立一个附加的事实, 即那个极限映射 (对应于上面定理中的 g) 是线性的.

定理 7.7 完备度量空间的闭子集是完备的. 任何度量空间的完备子集是闭的.

证明 此证明留给读者. □

由定理 7.7 知道, 一个完备的度量空间是*绝对闭的*, 它的意思是说不管我们如何把 A 扩张到一个更大的度量空间 B, A 总是 B 的闭子集. 实际上, 这个性质等价于完备性, 因为如果 A 不是完备的, 那么度量空间理论中一个非常重要的构造表明 A 可以被完备化. 那就是说, 我们可以构造一个包含 A 的完备度量空间 B. 现在如果 A 不完备, 则 A 在 B 中的闭包不同于 A 且是完备的, 因而 A 不是绝对闭的.

对一个度量空间的完备化构造可参看习题 7.21~7.23. 一个赋范线性空间的完备化当然是个巴拿赫空间.

定理 7.8 在定理 7.5 的背景下, 设 A 为度量空间, $\mathfrak{C}([A, W])$ 为从 A 到 W 的连续函数的空间, 并令

$$\mathfrak{BC}(A, W) = \mathfrak{B}(A, W) \cap \mathfrak{C}(A, W).$$

于是 $\mathfrak{BC}$ 是 $\mathfrak{B}$ 的闭子集.

证明 假设 $\{f_n\} \subset \mathfrak{BC}$ 且 $\|f_n - g\|_\infty \to 0$, $g \in \mathfrak{B}$. 我们必须证明 g 是连续的. 这是经常使用的 "向上, 在上, 向下"("up, over, and down") 论证法的应用, 它图示地被标出在图 4.3 中.

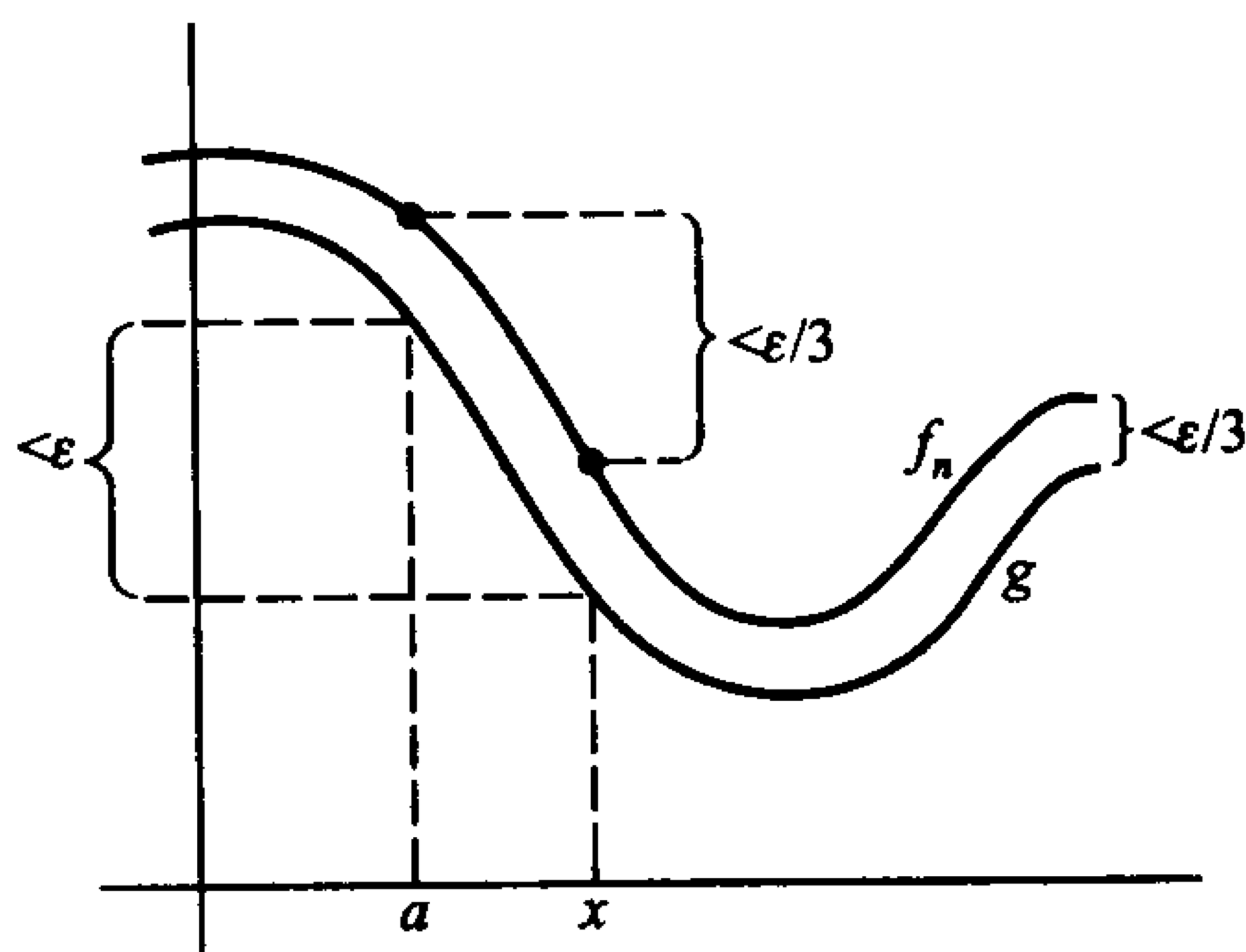


图 4.3

给出 ε, 我们首先选任意使 $\|f_n - g\|_\infty < \varepsilon/3$ 的 n. 现考虑任意 $a \in A$. 因为 f_n 在 a 连续, 故存在一个 δ 使得

$$\rho(x,a) < \delta \ \Rightarrow\ \|f_n(x) - f_n(a)\| < \varepsilon/3.$$

于是

$$\rho(x,a) < \delta \Rightarrow \|g(x) - g(a)\| \leqslant \|g(x) - f_n(x)\| + \|f_n(x) - f_n(a)\| \\ + \|f_n(a) - g(a)\| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

因此对每个 $a \in A, g$ 为连续, 从而 $g \in \mathfrak{BC}$ □

这个重要的经典结果习惯上是这样叙述的:连续函数的一致收敛序列的极限是连续的.

注 上面的证明要更广一点. 我们实际证明的是如果 $f_n \to f$ 是一致的, 并且如果每个 f_n 在 a 连续, 则 f 在 a 连续.

推论 7.4 $\mathfrak{BC}(A, W)$ 为巴拿赫空间.

定理 7.9 如果 A 为列紧度量空间, 则 A 为完备.

证明 A 中一个柯西序列有一个收敛于 A 中一个极限的子序列, 因此由引理 7.2 它自己也收敛于此极限. 故 A 为完备. □

在第 5 节中我们证明过一个紧集也是个完全有界集. 相反地可以证明一个完备的, 完全有界的集合 A 是列紧的, 故这两个性质合在一起等价于紧性.

关键的事实是如果 A 是完全有界的, 则 A 中的每个序列有一个柯西子序列. 如果 A 为完备, 此柯西子序列将收敛于 A 的一个点. 因此完全有界性和完备性合起来等价于紧性这个事实直接从下面的引理得到.

引理 7.4 *如果 A 为完全有界, 则 A 中每个序列都有一个柯西子序列.*

证明 设 $\{p_m\}$ 为 A 中任一序列. 由于 A 可由有限个半径 1 的球覆盖, 这样的球中至少有一个含有无穷多个 $\{p_m\}$ 中的点. 更准确地说, 存在一个无限集 $M_1 \subset \mathbb{Z}^+$ 使得集合 $\{p_m : m \in M_1\}$ 在单个的半径 1 的球内. 假设 $M_1, \cdots, M_n \subset \mathbb{Z}^+$ 已被定义使得 $M_{i+1} \subset M_i, i = 1, \cdots, n-1$, 且 M_n 为无限集. 而 $\{p_m : m \in M_i\}$ 是一个半径 $1/i$ 的球的子集, $i = 1, \cdots, n$. 由于 A 被有限个族半径为 $1/n+1$ 的球覆盖, 故至少有一个去覆盖的球含有 $\{p_m : m \in M_n\}$ 中的无穷多个点. 更准确地说, 存在一个无限集 $M_{n+1} \subset M_n$ 使得 $\{p_m : m \in M_{n+1}\}$ 是半径 $1/(n+1)$ 的球的一个子集. 我们因而定义了一个无穷序列 $\{M_n\}$, 其中每个 M_n 是 $\mathbb{Z}^+$ 的子集并具有上面的性质.

现选 $m_1 \in M_1$, $m_2 \in M_2$ 使 $m_2 > m_1$, 并且更一般地, 选 $m_{n+1} \in M_{n+1}$ 使 $m_{n+1} > m_n$. 那么子序列 $\{p_{m_n}\}_n$ 是柯西的. 因为给出 ε, 我们可选 n 使 $1/n < \varepsilon/2$. 于是 $i, j > n \Rightarrow m_i, m_j \in M_n \Rightarrow \rho(p_{m_i}, p_{m_j}) < 2(1/n) < \varepsilon$. 这证明了引理, 而我们的定理是个推论. □

定理 7.10 *一个度量空间 S 是列紧的当且仅当 S 为完全有界和完备的.*

接下来的三节将致力于完备性对微积分的应用, 但在着手这些极其重要的事之前我们应该讲几句关于无穷级数的话. 像在普通微积分那样, 如果 $\{\xi_n\}$ 是在一赋范线性空间 V 中的一个序列, 我们说级数 $\sum \xi_i$ 收敛并有和 α, 且记为 $\sum_1^\infty \xi_i = \alpha$ 指的是如果*部分和的序列收敛于*α. 这表明当 $n \to \infty$ 时 $\sigma_n \to \alpha$, 其中对每个 n, σ_n 等于有限和 $\sum_1^n \xi_i$. 我们说 $\sum \xi_i$*绝对收敛*是指如果范数的级数 $\sum \|\xi_i\|$ 在 $\mathbb{R}$ 中收敛. 除非每个绝对收敛的级数真的收敛, 否则这是在滥用术语, 而这个概念的重要性来自下面的定理.

定理 7.11 如果 V 是个巴拿赫空间, 则 V 中每个绝对收敛的级数是收敛的.

证明 设 $\sum \xi_i$ 绝对收敛. 这意味着 $\sum \|\xi_i\|$ 在 $\mathbb{R}$ 中收敛, 即序列 $\{s_n\}$ 在 $\mathbb{R}$ 中收敛, 其中 $s_n = \sum_1^n \|\xi_i\|$. 如果 $m < n$, 则

$$\|\sigma_n - \sigma_m\| = \left\| \sum_{m+1}^{n} \xi_i \right\| \leqslant \sum_{m+1}^{n} \|\xi_i\| = s_n - s_m.$$

因为 $\{s_i\}$ 在 $\mathbb{R}$ 中为柯西序列, 此不等式表明 $\{\sigma_i\}$ 在 V 中也是柯西序列, 并且因为 V 是完备的, 从而 $\{\sigma_n\}$ 在 V 中收敛. 那就是说 $\sum \xi_i$ 在 V 中收敛. □

读者被要求在一个习题中证明, 相反地, 如果一个赋范线性空间 V 满足每个绝对收敛的级数收敛, 则 V 为完备. 这个性质因而描绘了巴拿赫空间的特征.

我们将常常使用上面这个定理. 目前我们仅说明一个推论, 即经典的魏尔斯特拉斯比较判别法.

推论 7.5 如果 $\{f_n\}$ 是具有公共定义域的有界实 (或 W - 值的, 其中 W 为某个巴拿赫空间) 函数的序列, 并且如果存在一个正常数的序列 $\{M_n\}$ 使得 $\sum M_n$ 收敛和 $\|f_n\|_\infty \leqslant M_n$ 对每个 n 成立, 则 $\sum f_n$ 一致收敛.

证明 此假定蕴涵了 $\sum \|f_n\|_\infty$ 收敛, 故由定理, 在巴拿赫空间 $\mathfrak{B}(A, W)$ 中 $\sum f_n$ 收敛. 但 $\mathfrak{B}(A, W)$ 中的收敛性是一致收敛性. □

习题

7.1 证明度量空间中的柯西序列是个有界集.

7.2 设 V 是赋范线性空间. 证明 V 中两个柯西序列的和仍是柯西的.

7.3 证明如果 $\{\xi_n\}$ 在 V 中为柯西序列, $\{a_n\}$ 为 $\mathbb{R}$ 中的柯西序列, 则 $\{a_n\xi_n\}$ 也是 V 中的柯西序列.

7.4 证明如果 $\{\xi_n\}$ 是在一个赋范线性空间 V 中的柯西序列, 则 $\{\|\xi_n\|\}$ 是 $\mathbb{R}$ 中的柯西序列..

7.5 证明如果 $\{x_n\}$ 和 $\{y_n\}$ 为两个在度量空间 S 中的柯西序列, 则 $\{\rho(x_n, y_n)\}$ 是 $\mathbb{R}$ 中的柯西序列.

7.6 证明在引理 7.2 的证明后面所作的那段陈述.

7.7 有理数系是一个不完备的度量空间. 展示一个有理数的柯西序列它不收敛于一个有理数以证明此论断.

7.8 证明定理 7.1.

7.9 由定理 7.1 推导出定理 7.3 的一个加强形式.

7.10 模仿定理 7.5 的证明, 写出定理 7.6 的一个仔细的证明.

7.11 证明定理 7.7.

7.12 设度量空间 X 有一个稠子集 Y 使得 Y 中的每个柯西序列在 X 中收敛. 证明 X 为完备.

7.13 证明在一个赋范线性空间 V 中所有柯西序列的集合 W 自身也是个向量空间, 并在 W 上可用 $p(\{\xi_n\}) = \lim\|\xi_n\|$ 定义一个半范数 p.(由正文中的素材和前面的问题将此拼凑出来.)

7.14 继续上面的习题, 对每个 $\xi \in V$ 让 ξ^c 为所有项均是 ξ 的常值序列. 证明 $\theta : \xi \mapsto \xi^c$ 是 V 到 W 内的一个等距线性单射, 并且 $\theta[V]$ 在上面习题中的半范数下稠于 W.

7.15 下一步证明在 $\theta[V]$ 中的每个柯西序列在 W 中收敛. 将第三章的习题 4.18 和本章的习题 7.12~7.14 集中一起得出结论说, 如果 N 是 W 中的零柯西序列的集合, 则 W/N 是个巴拿赫空间, 而 $\xi \mapsto \xi^c$ 是从 V 到 W/N 的一个稠子空间的一个等距线性单射. 这构成赋范线性空间 V 的标准完备化.

7.16 我们现在描述一个构造度量空间 S 的完备化的非标准的方法. 选取 S 中某个点 p_0, 并设 V 为 S 上所有满足 $f(p_0) = 0$ 和 f 为利普希茨函数的所有实函数 f 的集合. 对 $f \in V$ 定义 $\|f\|$ 为 f 的最小利普希茨常数, 即

$$\|f\| = \operatorname*{lub}_{p \neq 0}\{|f(p) - f(q)|/\rho(p,q)\}.$$

证明在此范数下 V 是个赋范线性空间.(V 实际是完备的, 但我们不需要用它.]

7.17 继续上面习题. 我们知道, 由定理 7.6V 上的所有有界线性泛函的空间即对偶空间 V^* 是完备的. 我们现在要证明 S 可以等距地嵌入到 V^* 中; 于是作为 V 的子集 S 的闭包将是所想要的 S 的完备化. 对每个 $p \in S$, 设 $\theta_p : V \to \mathbb{R}$ 为 "在 p 的赋值". 那就是说 $\theta_p(f) = f(p)$. 证明 $\theta_p \in V^*$ 且 $\|\theta_p - \theta_q\| \leqslant \rho(p,q)$.

7.18 为了得出映射 $\theta : p \mapsto \theta_p$ 是个等距映射的结论 (即保距的), 我们必须证明相反的不等式 $\|\theta_p - \theta_q\| \geqslant \rho(p,q)$. 要做到这点, 选取 p 并考虑特定的函数 $f(x) = \rho(p,x) - \rho(p,p_0)$. 证明 f 在 V 中且 $\|f\| = 1$(由本章早先的引理得出). 现在应用 $\|\theta_p - \theta_q\|$ 的定义并最后有结论说 θ 是一个从 S 到 V^* 内的等距单射. 于是 $\overline{\theta[s]}$ 是我们所构造的完备化.

7.19 证明如果一个赋范线性空间 V 有性质即每个绝对收敛的级数收敛, 则 V 为完备.(设 $\{\alpha_n\}$ 为一柯西序列. 证明有一子序列 $\{\alpha_{n_i}\}_i$ 使得如果 $\xi_i = \alpha_{n_{i+1}} - \alpha_{n_i}$, 则 $\|\xi_i\| < 2^{-i}$. 结论是子序列收敛, 从而完成证明.)

7.20 上面的习题对 V 为完备性给出了一个非常有用的制定准则. 用它证明如果 V 为巴拿赫空间, N 为闭子空间, 则 V/N 为巴拿赫空间 (对 V/N 上的范数见第三章的习题 4.14).

7.21 证明无穷小量的一致收敛级数 (这些无穷小量有相同的定义域) 的和是个无穷小量.

4.8 巴拿赫代数初探

当我们在上一章在考虑隐函数定理和反函数定理时, 我们看到如果知道了一个变换 T 有一逆 T^{-1}, 则只要 $\|S-T\|$ 足够小 S 也有逆, 以及映射 $T \mapsto T^{-1}$ 在所有可逆元组成的开集上为连续这样的事实是多么地有用. 当所考虑的空间是有限维时, 这些事实可以由从 $\mathrm{Hom}V$ 到 $\mathbb{R}$ 的行列式函数 $T \mapsto \Delta(T)$ 的连续性得出. 也有可能用 T 的上界和下界以及它的紧密近似 S 来直接论证从而得出这些事实. 但是最自然, 最优美以及 (在巴拿赫空间情形中) 最简单的用来证明这些事实的方式是证明, 如果 V 是巴拿赫空间且 T 在 $\mathrm{Hom}V$ 中有小于 1 的范数, 则几何级数 $\sum_0^\infty T^n$ 的和是 $I-T$ 的逆, 这恰如在初等微积分中那样. 但在做这个论证中, T 是个线性变换这个事实没有起什么重要作用, 我们暂且离题去探索一下这个情形.

让我们概括一下当 V 是巴拿赫空间时 $\mathrm{Hom}V$ 的范数与代数性质. 首先, 我们知道 $\mathrm{Hom}V$ 也是巴拿赫空间, 其次, 它是个代数. 那就是说, 它具有一个可结合的乘法运算 (复合), 此运算按下述法则与线性运算相关联:

$$S(T_1+T_2)=ST_1+ST_2,$$

$$(S_1+S_2)T=S_1T+S_2T,$$

$$c(ST)=(cS)T=S(cT).$$

最后, 乘法与范数的关联是

$$\|ST\| \leqslant \|S\|\,\|T\| \quad 和 \quad \|I\|=1.$$

这一系列的性质完全构成了对巴拿赫代数的公理.

对于函数的某些性质我们可以用忘掉它们是函数而只把它们想成是一个向量空间的元素的办法, 便能把这些性质看得最清楚; 恰如这种情况一样, 现在也可以最终证明, 对于在 $\mathrm{Hom}V$ 中变换的某些性质当我们忘掉线性变换的复杂性

质仅仅把它考虑为一个抽象的巴拿赫代数 A 中一个元素时我们可以最简单地处理这些性质.

我们可以在一个巴拿赫代数中做而不能在巴拿赫空间做的最重要的简单事情便是考虑幂级数. 下面的定理表明, 几何级数像在初等微积分中那样, 在这里特别地起着一个中心的作用. 由于我们不把 A 中的元素想成为变换, 我们将以小写的字母表示它们;e 是 A 中的单位元.

定理 8.1 如果 A 是个巴拿赫代数, 并且如果 x 在 A 中的范数小于 1, 则 $e-x$ 可逆, 且其逆为用 x 表示的几何级数的和:

$$(e-x)^{-1}=\sum_0^{\infty}x^n.$$

又, $\|e-(e-x)^{-1}\|\leqslant r/(1-r)$, 其中 $r=\|x\|$.

证明 因为 $\|x^n\|\leqslant\|x\|^n=r^n$, 所以级数 $\sum x^n$ 当 $\|x\|<1$ 时绝对收敛, 这只要与普通的几何级数 $\sum r^n$ 比较便知. 因此它收敛, 且如果 $y=\sum_0^{\infty}x^n$, 则

$$(e-x)y=\lim_{n\to\infty}(e-x)\sum_0^n x^i=\lim(e-x^{n+1})=e,$$

这里由于有 $\|x\|^{n+1}\leqslant r^{n+1}\to 0$ 的缘故. 此即 $y=(e-x)^{-1}$. 最后,

$$\|e-(e-x)^{-1}\|=\left\|\sum_1^{\infty}x^n\right\|\leqslant\sum_1^{\infty}r^n=r/(1-r).$$

□

定理 8.2 在巴拿赫代数 A 中可逆元的集合 n 为开集且映射 $x\to x^{-1}$ 为 n 到 n 的连续映射. 事实上, 如果 y^{-1} 存在, $m=\|y^{-1}\|$, 则 $(y-h)^{-1}$ 只要 $\|h\|<1/m$ 便存在, 而且 $\|(y-h)^{-1}-y^{-1}\|\leqslant m^2\|h\|/(1-m\|h\|)$.

证明 令 $x=y^{-1}h$. 于是 $(y-h)=y(e-x)$, 其中 $\|x\|=\|y^{-1}h\|\leqslant m\|h\|$, 故由前一个定理知 $y-h$ 可逆, 而且在 $\|h\|<1/m$ 假定下 $(y-h)^{-1}=(e-x)^{-1}y^{-1}$. 于是也有

$$\|y^{-1}-(y-h)^{-1}\|\leqslant\|e-(e-x)^{-1}\|\cdot m,$$

由上面定理中最后一个不等式知它被

$$mr/(1-r)\leqslant m^2\|h\|/(1-m\|h\|)$$

界于上.

□

推论 8.1 如果 V 和 W 为巴拿赫空间, 则 $\mathrm{Hom}(V,W)$ 中的可逆元构成一个开集, 且映射 $T \mapsto T^{-1}$ 在此区域上连续.

证明 假设 T^{-1} 存在, 令 $m = \|T^{-1}\|$. 于是如果 $\|T - S\| < 1/m$, 我们有 $\|I - I^{-1}S\| \leqslant \|T^{-1}\|\,\|T - S\| < 1$, 故 $T^{-1}S = I - (I - T^{-1}S)$ 是 $\mathrm{Hom}V$ 中的可逆元. 因此 $S = T(T^{-1}S)$ 可逆且 $S^{-1} = (T^{-1}S)^{-1}T^{-1}$. 至于 $S \mapsto S^{-1}$ 的连续性留给读者去证明. □

我们在上面看到从巴拿赫代数 A 中的单位开球 $B_1(0)$ 到 A 的映射 $x \mapsto (e-x)^{-1}$ 是几何幂级数的和. 我们还可以用收敛幂级数去定义许多其他的映射而几乎不需太费劲.

定理 8.3 设 A 为巴拿赫代数. 设序列 $\{a_n\} \subset A$ 和正数 δ 使得序列 $\{\|a_n\|\delta^n\}$ 为有界. 于是 $\sum a_n x^n$ 在 A 中的球 $B_\delta(0)$ 中对 x 收敛, 并且如果 $0 < s < \delta$, 则此级数在 $B_s(0)$ 上一致收敛.

证明 令 $r = s/\delta$, 并设 b 为序列 $\{\|a_n\|\delta^n\}$ 的一个界. 在球 $B_s(0)$ 上我们有 $\|a_n x^n\| \leqslant \|a_n\|s^n = \|a_n\|\delta^n r^n \leqslant br^n$, 因而此级数用几何级数 $b\sum r^n, r < 1$ 比较, 知其在此球上一致收敛. □

我们最感兴趣的级数都具有实系数. 由于向量 x 和纯量 t 的积是代数积 $(te)x$, 故这些级数在我们上面的讨论之中. 除了处理上面的几何级数外, 我们特别对指数函数$e^x = \sum_0^\infty x^n/n!$ 感兴趣. 初等微积分中通常的比较论证表明在这里一样容易得出此级数对 A 中每个 x 均收敛而且在任意的球上一致收敛.

自然要考虑这些收敛级数定义的从 A 到 A 的映射的可微性, 我们从关于一个序列极限的可微性的基本定理开始, 在下面陈述这些基础性的东西.

定理 8.4 设 $\{F^n\}$ 为赋范线性空间 V 中一个球 B 到赋范空间 W 的映射的序列, 使得 F^n 在 B 上逐点收敛于映射 F, 并使得 $\{dF_\alpha^m\}$ 对每个 α 收敛并于 α 是一致的. 则 F 在 B 上可微, 而且 $dF_\beta = \lim F_\beta^n$ 对 B 中每个 β 成立.

证明 固定 β 并令 $T = \lim F_\beta^n$. 由 $\{dF^n\}$ 的一致收敛性, 对给出的 ε, 存在一个 N 使得 $\|dF_\alpha^n - dF_\alpha^N\| \leqslant \varepsilon$ 对所有 $n \geqslant N$ 和所有 $\alpha \in B$ 成立. 于是从微分的中值定理得到

$$\|(\Delta F_\beta^n(\xi) - \Delta F_\beta^N(\xi)) - (dF_\beta^n(\xi) - dF_\beta^N(\xi))\| \leqslant 2\varepsilon\|\xi\|$$

对所有 $n \geqslant N$ 和所有使 $\beta+\xi \in B$ 的 ξ 成立. 让 $n \to \infty$ 并重新编组, 我们有

$$\|(\Delta F_\beta(\xi) - T(\xi)) - (\Delta F_\beta^N(\xi) - dF_\beta^N(\xi))\| \leqslant 2\varepsilon\|\xi\|$$

对所有这样的 ξ 成立. 但是由 dF_β^N 的定义, 存在一个 δ 使得当 $\|\xi\| < \delta$ 有

$$\|\Delta F_\beta^N(\xi) - dF_\beta^N(\xi)\| \leqslant \varepsilon\|\xi\|.$$

将最后的这两个不等式合并, 我们看出

$$\|\xi\| < \delta \ \Rightarrow \ \|\Delta F_\beta(\xi) - T(\xi)\| \leqslant 3\varepsilon\|\xi\|.$$

因此 F 在 β 可微且 $dF_\beta = T$. □

后面留下的证明被当作一组习题.

引理 8.1 巴拿赫代数 A 上的乘法是可微的 (从 $A \times A$ 到 A). 如果设 p 为此乘法函数, 则 $p(x, y) = xy$, 于是

$$dp_{\prec a,b\succ}(x, y) = ay + xb.$$

引理 8.2 设 A 是一个交换的巴拿赫代数, 并设 p 是单项式函数 $p(x) = ax^n$. 且 p 处处可微且 $dp_y(x) = nay^{n-1}x$.

引理 8.3 如果 $\{\|a_n\|r^n\}$ 是 $\mathbb{R}$ 中的有界序列, 则 $\{n\|a_n\|s^n\}$ 对 $0 < s < r$ 有界, 因而 $\sum na_nx^{n-1}$ 在 A 中比球 $B_r(0)$ 更小的任意球上一致收敛.

定理 8.5 如果 A 是个交换的巴拿赫代数, 且 $\{a_n\} \subset A$ 使 $\{\|a_n\|r^n\}$ 在 $\mathbb{R}$ 中有界, 则 $F(x) = \sum_0^\infty a_nx^n$ 在 A 中球 $B_r(0)$ 上有定义并可微, 而且

$$dF_y(x) = \left(\sum_1^\infty na_ny^{n-1}\right) \cdot x.$$

自然称元 $\sum_1^\infty na_ny^{n-1}$ 为 F 在 y 的导数并以 $F'(y)$ 表示, 虽然它不合于我们的规则即导数作为差商的极限所得到的向量. 上面定理的值得注意的方面是, 对于此类从 A 的开子集到 A 的可微映射而言, 线性变换 dF_y 是乘以 A 中的一个元:$dF_y(x) = F'(y) \cdot x$.

特别, 指数函数 $\exp(x) = e^x = \sum_0^\infty x^n/n!$ 是自己的导数, 因为 $\sum_1^\infty nx^{n-1}/n! = \sum_0^\infty x^m/m!$, 并且由这个事实 (见习题) 或从对所考虑的这个级数的直接代数

处理, 我们可以推导出指数法则 $e^{x+y}=e^xe^y$. 但是请记住, 这是在交换巴拿赫代数上. 函数 $x\mapsto e^x=\sum_0^\infty x^n/n!$ 可以在任何巴拿赫代数 A 上被容易地定义, 然而当 A 是非交换时它就几乎不是那么令人愉快了. 但是有一件我们总是能做的, 并常常因此而扭转了局面的事, 即把此指数映射限制到 A 的一个可交换的子代数上, 譬如由单个元素 x 生成的代数. 举例来说, 我们可以考虑在任一巴拿赫代数 A 内的参数弧 $\gamma(t)=e^{tx}$(x 固定), 并且因为它的值域位于由 x 生成的交换子代数 X 中, 我们便能应用第三章的定理 7.2 得出 γ 是可微的结论, 并且有

$$\gamma'(t)=\mathrm{dexp}_{tx}(x)=xe^{tx}.$$

这也可直接从指数法则容易地证得:

$$\Delta\gamma_t(h)=e^{(t+h)x}-e^{tx}=e^{tx}(e^{hx}-1),$$

并因为从这个级数中显见有 $(e^{hx}-1)/h$ 当 $h\to 0$ 时趋向 x, 从而有

$$\gamma'(t)=\lim_{h\to 0}\frac{\Delta\gamma_t(h)}{h}=xe^{tx}.$$

习题

8.1 完成定理 8.2 的推论的证明.

8.2 设 A 为巴拿赫代数, $\{a_n\}\subset R$ 及 $x\in A$ 使得 $\sum a_ix^i$ 收敛. 假设 x 也满足一个多项式等式 $p(x)=\sum_0^n b_ix^i=0$, 其中 $\{b_i\}\subset\mathbb{R}$ 且 $b_n\neq 0$. 证明元素 $\sum_1^\infty a_ix^i$ 是一次数 $\leqslant n-1$ 的 x 的多项式.(设 M 为 $\{x^i\}_0^{n-1}$ 的线性张成空间, 首先证明对所有 i, $x^i\in M$.)

8.3 设为任一巴拿赫代数, x 为 A 中一个固定元, 又设 X 是包含 x 的 A 中最小的闭子代数. 证明 X 为交换巴拿赫代数. (多项式 $p(x)=\sum_0^n a_ix^i$ 的集合是包含 x 的最小代数. 考虑它在 A 中的闭包.)

8.4 证明引理 8.1.[提示:$\prec x,y\succ\mapsto xy$ 是个有界双线性映射.]

8.5 用直接的 Δ 估算, 像在初等微积分中那样从二项式展开证明引理 8.2.

8.6 由引理 8.1 用归纳法证明引理 8.2.

8.7 设 A 为任一巴拿赫代数. 证明 $p:x\mapsto x^3$ 可微并且 $dp_a(x)=xa^2+axa+a^2x$.

8.8 用归纳法证明如果 $q(x)=x^n$, 则 q 是可微的并且

$$dg_a(x)=\sum_{i=0}^{n-1}a^ixa^{(n-1-i)}.$$

作为其推论导出引理 8.2.

8.9 设 A 为任一巴拿赫代数. 证明 $r: x \to x^{-1}$ 在可逆元组成的开集 U 上处处可微, 并且

$$dr_a(x) = -a^{-1}xa^{-1}.$$

[提示: 考查定理 8.1 和 8.2 的证明.]

8.10 设 A 为一赋范线性空间 V 的开子集, F 和 G 为从 A 到一个巴拿赫代数X 的映射, 它们在 α 可微. 证明乘积映射 FG 在 α 可微且 $d(FG)_\alpha = F(\alpha)dG_\alpha + dF_\alpha G(\alpha)$. 它可以推出 $d(F^2)_\alpha = 2F(\alpha)dF_\alpha$ 吗?

8.11 继续上一个习题, 证明如果 X 为交换巴拿赫代数, 则 $d(F^n)_\alpha = nF^{n-1}(\alpha)dF_\alpha$.

8.12 设 $F: A \to X$ 为从一赋范线性空间中开集 A 到一个巴拿赫代数 X 的可微映射, 并假设元素 $F(\xi)$ 在 X 中可逆, 其中 ξ 为 A 中每个向量. 证明映射 $G: \xi \mapsto [F(\xi)]^{-1}$ 可微, 且 $dG_\alpha(\xi) = -F(\alpha)^{-1}dF_\alpha(\xi)F(\alpha)$. 再证明如果 F 是个参数弧 $(A = I \subset \mathbb{R})$, 则 $G'(\alpha) = -F(\alpha)^{-1}F'(\alpha)F(\alpha)^{-1}$.

8.13 证明引理 8.3.

8.14 指出引理 8.3 使得定理 8.4 变得适合于应用, 并以此证明定理 8.5.

8.15 证明在定理 8.4 中, 假设我们已知上域空间 W 是个巴拿赫空间则 $\{F^n\}$ 到 F 的收敛性只须假定在一个点上就可以了.

8.16 我们需要证明对在一个交换巴拿赫代数上的指数函数的指数法则. 先证明 $(\exp(-x))\cdot(\exp x) \equiv e$, 还需应用第三章的习题 7.13, 上面的习题 8.10 和 $d\exp_a(x) = \exp(a)x$.

8.17 证明如果 X 为交换巴拿赫代数, $F: X \to X$ 为可微映射使得 $dF_\alpha(\xi) = \xi F(\alpha)$, 则有某个常数 β 使 $F(\xi) = \beta\exp\xi$. [考虑 $F(\xi)\exp(-\xi)$ 的微分.]

8.18 现令 $F(\xi) = \exp(\xi + \eta)$ 并证明

$$\exp(\xi + \eta) = \exp(\xi)\exp(\eta),$$

这可由上面的习题得到. 你还需要 $\exp 0 = 1$.

8.19 设 z 为一交换巴拿赫代数 X 中的一个幂零元. 就是说, 对某个正整数 p 有 $z^p = 0$. 用基于二项展开式的一个初等估值证明, 如果 $\|x\| < 1$, 则 $\|x + z\|^n \leqslant kn^p\|x\|^{n-p}$ 对 $n > p$ 成立. 正项级数 $\sum n^a r^n$ 对 $r < 1$ 收敛 (由比率判别法). 因此证明当 $\|x\| < 1$ 时对 $\log(1 - (x + z))$ 的级数和对 $(1 - (x + z))^{-1}$ 的级数均收敛.

8.20 继续上面的习题, 证明 $F(y) = \log(1 - y)$ 在球 $\|y - z\| < 1$ 上有定义并可微以及 $dF_a(x) = -(1 - a)^{-1} \cdot x$. 因此证明 $\exp(\log(1 - y)) = 1 - y$ 在此球上成立, 这要么用逆映射定理要么用微分的复合函数规则. 最后得出对每个 X 的幂零元 z, 存在 X 中一个 u, 使得 $\exp u = 1 - x$ 的结论.

8.21 设 $X_1, \cdots, X_n$ 为巴拿赫代数. 证明在积巴拿赫空间 $X = \prod_1^n X_i$ 中如果乘积 $\boldsymbol{xy} = \prec x_1, \cdots, x_n \succ \prec y_1, \cdots, y_n \succ$ 定义为 $\prec x_1 y_1, \cdots, x_n y_n \succ$, 并在 X 在采用最大值范数, 则 X 成为一个巴拿赫代数.

8.22 在上面的情形中, 投影 π_i 现在成为了有界代数同态. 事实上, 恰似在积空间上我们原来的向量定义那样, 我们在 X 上的乘法定义被要求 $\pi_i(\boldsymbol{xy}) = \pi_i(\boldsymbol{x})\pi_i(\boldsymbol{y})$ 对所有 i 成立所确定. 叙述并证明一个类比于第一章的定理 3.4 的代数定理.

8.23 继续上面的讨论, 假设级数 $\sum \boldsymbol{a}_n \boldsymbol{x}^n$ 在 X 中收敛于和 $\boldsymbol{y}$. 证明 $\sum (a_n)_i (x_i)^n$ 对每个 i 在 X_i 中收敛于 y_i, 其中自然 $\boldsymbol{y} = \prec y_1, \cdots, y_n \succ$. 最后得出 $e^{\boldsymbol{x}} = \prec e^{x_1}, \cdots, e^{x_n} \succ$ 对 X 中任意 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ$ 成立.

8.24 在一个交换巴拿赫代数上定义正弦和余弦函数, 并证明 $\sin' = \cos, \cos' = -\sin, \sin^2 + \cos^2 = e$.

4.9 压缩映射不动点定理

在本节我们将对压缩映射证明一个非常简洁和优美的不动点定理, 然后再运用它去完成隐函数定理的证明. 后面, 在第六章中, 它将是我们证明常微分方程的基本的存在性和惟一性定理的基础. 本节最后以比较不动点的迭代过程和牛顿法的迭代过程作为结束.

从一个度量空间 X 到自己的映射 K 是一个压缩是说它是一个利普希茨映射, 其常数小于 1; 那就是说, 如果存在一个常数 C 满足 $0 < C < 1$ 使得 $\rho(K(x), K(y)) \leqslant C\rho(x, y)$ 对所有 $x, y \in X$ 成立. 当然, K 的一个不动点是指一个点 x 使 $K(x) = x$.

一个压缩 K 最多只能有一个不动点, 因为如果 $K(x) = x$ 和 $K(y) = y$, 则 $\rho(x, y) = \rho(K(x), K(y)) \leqslant C\rho(x, y)$, 因而 $(1 - C)\rho(x, y) \leqslant 0$. 但 $C < 1$, 这表明 $\rho(x, y) = 0$, 从而 $x = y$.

定理 9.1 设 X 为一个非空的完备度量空间, 并设 $K : X \to X$ 为一个压缩. 则 K 有一个 (惟一的) 不动点.

证明 在 X 中选取任意一个点 x_0, 并归纳地定义序列 $\{x_n\}_0^\infty$ 为 $x_1 = K(x_0)$, $x_2 = K(x_1) = K^2(x_0)$, 和 $x_n = K(x_{n-1}) = K^n(x_0)$. 令 $\delta = \rho(x_1, x_0)$. 于是

$\rho(x_2, x_1) = \rho(K(x_1), K(x_0)) \leqslant C\rho(x_1, x_0) = C\delta$, 然后由归纳法,

$$\rho(x_{n+1}, x_n) = \rho(K(x_n), K(x_{n-1})) \leqslant C\rho(x_n, x_{n-1}) \leqslant C \cdot C^{n-1}\delta$$
$$= C^n\delta.$$

因此得出 $\{x_n\}$ 是柯西的, 因为如果 $m > n$, 则

$$\rho(x_m, x_n) \leqslant \sum_n^{m-1} \rho(x_{i+1}, x_i) \leqslant \sum_n^{m-1} C^i\delta < C^n\delta/(1-C),$$

而当 $n \to \infty$ 时 $C^n \to 0$, 这是因为 $C < 1$. 由于 X 为完备, 故 $\{x_n\}$ 收敛于 X 中某个 a, 从而得到 $K(a) = \lim K(x_n) = \lim x_{n+1} = a$, 故 a 是个不动点. □

在实践中, 我们所遇到的映射 K 只是在某个特定点 p 附近是压缩的, 我们不得不设立一个 p 的适当邻域被 K 带进它自身. 我们在下面证明如果 K 在一个 p 的球上是个压缩, 并且如果 K 不把中心 p 移动得非常远, 则可应用这个定理.

推论 9.1 设 D 是一个完备度量空间 X 中的一个闭球, 并让 $K: D \to X$ 为一个压缩, 它把 D 的中心移动的距离最多为 $(1-C)r$, 其中 r 是 D 的半径, C 为压缩常数. 则 K 有一个惟一的不动点且它在 D 中.

证明 我们仅需检验 K 的值域实际在 D 中. 如果 p 是 D 的中心, x 为 D 中任一点, 于是

$$\rho(K(x), p) \leqslant \rho(K(x), K(p)) + \rho(K(p), p)$$
$$\leqslant C\rho(x, p) + (1-C)r \leqslant Cr + (1-C)r = r.$$ □

推论 9.2 设 B 为完备度量空间 X 的一个开球, 并让 $K: B \to X$ 为一个压缩, 它移动 B 的中心的距离小于 $(1-C)r$, 其中 r 为 B 的半径, C 为压缩常数. 于是 K 有一个惟一的不动点.

证明 将 K 限制在任意稍稍小一点的与 B 同心的一个闭球上, 然后用前面那个推论. □

推论 9.3 设 K 为完备度量空间上的一个压缩, 并假设 K 移动点 x 的距离为 d. 于是从 x 到此不动点的距离最多为 $d/(1-C)$, 这里的 C 为压缩常数.

证明 设 D 为 x 的一个闭球, 其半径为 $r = d/(1-C)$, 并应用推论 1 于 K 在 D 上的限制. 它表明此不动点在 D 中. □

我们现在假定压缩 K 包含了一个参数 s, 故 K 现在是两个变量的函数 $K(s,x)$, 我们将假设 K 是一个在 s 上一致的对 x 的压缩, 它的意思是 $\rho(K(s,x), K(s,y)) \leqslant C\rho(x,y)$ 对所有 x,y 和 s 成立, 其中 $0<c<1$. 我们也将假定 K 对每个固定的 x 是 s 的连续函数.

推论 9.4 *设 K 为从 $S\times X$ 到 X 的一个映射, 其中 X 是个完备度量空间, S 为一个任意的度量空间, 并设 $K(s,x)$ 是在 x 为压缩而对 s 为一致, 又对每个 x 它是 s 的连续函数. 于是不动点 p_s 是 s 的连续函数.*

证明 给出 ε, 我们用 K 对第一个变量在点 $\prec t,p_t \succ$ 的连续性选择 δ, 使得结果 $\rho(s,t)<\delta$, 则由 $K(s,p_t)$ 到 $K(t,p_t)$ 的距离最多为 ε. 由于 $K(t,p_t)=p_t$, 这只不过是说具参数值 s 的压缩移动了 p_t 的距离最多为 ε, 因而由 p_t 到不动点 p_s 的距离按推论 9.3 最多为 $\varepsilon/(1-c)$. 那就是说, $\rho(s,t)<\delta \Rightarrow \rho(p_t,p_s)<\varepsilon/(1-C)$, 其中 C 为一致压缩常数, 映射 $s\mapsto p_s$ 因此在 t 为连续. □

联合推论 9.2 和 9.4 我们有下列定理.

定理 9.2 *设 B 为完备度量空间 X 中的一个球, S 为任一度量空间, 并设 K 为从 $S\times B$ 到 X 的映射, 它在第一个变量上一致地对第二个变量为一个压缩. 又假设 K 对每个 S 中的 s 移动了 B 的中心一个小于 $(1-C)r$ 的距离, 其中 r 是 B 的半径, C 是一致压缩常数. 于是对 S 中每个 s 存在一个惟一的 $p\in B$ 使得 $K(s,p)=p$, 并且映射 $s\mapsto p$ 从 S 到 B 是连续的.*

我们现在可以完成隐函数定理的证明了.

定理 9.3 *设 V,W 和 X 为巴拿赫空间, $A\times B$ 是 $V\times W$ 的开子集, $G: A\times B\to X$ 为连续并有连续的第二个偏微分. 假设 $\prec \alpha,\beta \succ$ 为 $A\times B$ 中使 $G(\alpha,\beta)=0$ 和 $dG^2_{\prec\alpha,\beta\succ}$ 可逆的点. 于是存在 α 和 β 的开球, 分别表为 M,N, 使得对每个 $\xi\in M$ 存在一个惟一的 $\eta\in N$ 满足 $G(\xi,\eta)=0$. 而在 $\prec\alpha,\beta\succ$ 附近被条件 $G(\xi,F(\xi))=0$ 惟一定义的函数 F 连续.*

证明 令 $T=dG^2_{\prec\alpha,\beta\succ}$, $K(\xi,\eta)=\eta - T^{-1}(G,(\xi,\eta))$. 于是 K 是从 $A\times B$ 到 W 的连续映射, 满足 $K(\alpha,\beta)=\beta$, 并且 K 有连续的第二个偏微分使得 $dK^2_{\prec\alpha,\beta\succ}=$

0. 因为 $dK^2_{\prec\mu,\nu\succ}$ 是 $\prec\mu,\nu\succ$ 的连续函数, 我们可选取 $\prec\alpha,\beta\succ$ 的一个球的乘积 $M\times N$, 在其上 $dK^2_{\prec\mu,\nu\succ}$ 以 1/2 有界, 并且如有必要可减小球 M 使得对 M 中的 μ 我们也有 $\|K(\mu,\beta)-\beta\|<r/2$, 其中 r 是球 N 的半径. 微分的中值定理表明 K 对它的第二个变量为具有常数 1/2 的压缩. 前面的定理因而表明对 M 中的每个 ξ 存在一个惟一的 $\eta\in N$ 使得 $K(\xi,\eta)=\eta$ 并且映射 $F:\xi\mapsto\eta$ 为连续. 因为 $K(\xi,\eta)=\eta$ 当且仅当 $G(\xi,\eta)=0$, 证完. □

定理 8.2 和 9.3 列出了隐函数定理的所有成分.(但可参看习题 9.8.)

我们下面从另一个角度指出, 如果一个依赖于参数的压缩映射为连续可微, 则其不动点是参数的连续可微函数.

定理 9.4 *设 V 和 W 为巴拿赫空间, K 为从 $V\times W$ 的一个开子集 $A\times B$ 到 W 的可微映射, 并满足定理 9.2 的假设条件. 则由方程 $K(\xi,F(\xi))=F(\xi)$ 惟一定义的从 A 到 B 的函数 F 是可微的.*

证明 不等式 $\|K(\xi,\eta')-K(\xi,\eta'')\|\leqslant C\|\eta'-\eta''\|$ 等价于对 $A\times B$ 中所有的 $\prec\alpha,\beta\succ$ 有 $\|dK^2_{\prec\alpha,\beta\succ}\|\leqslant C$. 现定义 G 为 $G(\xi,\eta)=\eta-K(\xi,\eta)$, 并看出 $dG^2=I-dK^2$, 从而 dG^2 由定理 8.1 为可逆. 因为 $G(\xi,F(\xi))=0$, 从而由第三章的定理 11.1 得出 F 为可微, 且其微分由微分前面的方程得到. □

推论 9.5 *如果 K 连续可微, 则 F 也如此.*

* 我们应该强调指出, 不动点定理不仅将隐函数定理作为其推论而且不动点定理的证明给出了实际求 $F(\xi)$ 值的一个迭代过程; 当然是在我们知道了如何计算 T^{-1}(这里的 $T=dG^2_{\prec\alpha,\beta\succ}$) 的条件下才能做到. 事实上, 对在 $\prec\alpha,\beta\succ$ 的一个足够小的球内给出的 ξ 值, 考虑函数 $G(\xi,\cdot)$. 如果我们令 $K(\xi,\eta)=\eta-T^{-1}(G(\xi,\eta))$, 则归纳过程

$$\eta_{i+1}=K(\xi,\eta_i)$$

成为

$$(\eta_{i+1}-\eta_i)=-T^{-1}G(\xi,\eta_i). \tag{9.1}$$

这个迭代过程的意思可通过研究当 $V=W=\mathbb{R}^1$ 时的图像很容易明白 (见图 4.4). 正如上面所证明的, 在适当的假设条件下, 级数 $\sum\|\eta_{i+1}-\eta_i\|$ 几何地收敛.

将这个过程与初等微积分中的牛顿法进行比较是颇有教益的. 在那里的迭代方案 (9.1) 被换为

$$(\eta_{i+1}-\eta_i)=S_i^{-1}G(\xi,\eta_i), \tag{9.2}$$

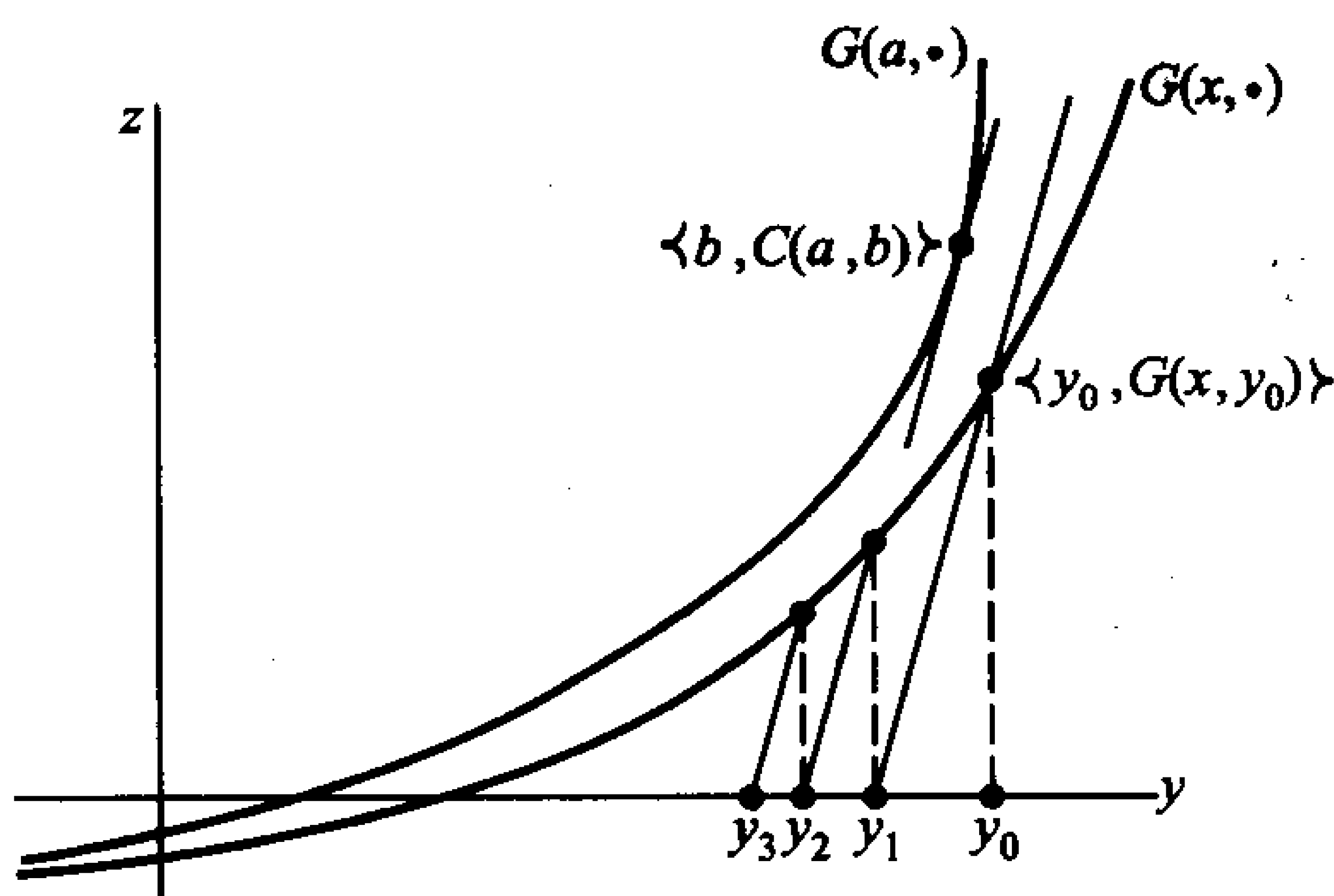


图 4.4

其中 $S_i = dG^2_{\prec\xi,\eta_i\succ}$(见图 4.5). 我们将看到, 这个过程 (当它可行时) 比 (9.1) 式的收敛快得多, 但它要受到必须无穷多次计算线性变换 S_i 的逆所带来的不利因素的折磨.

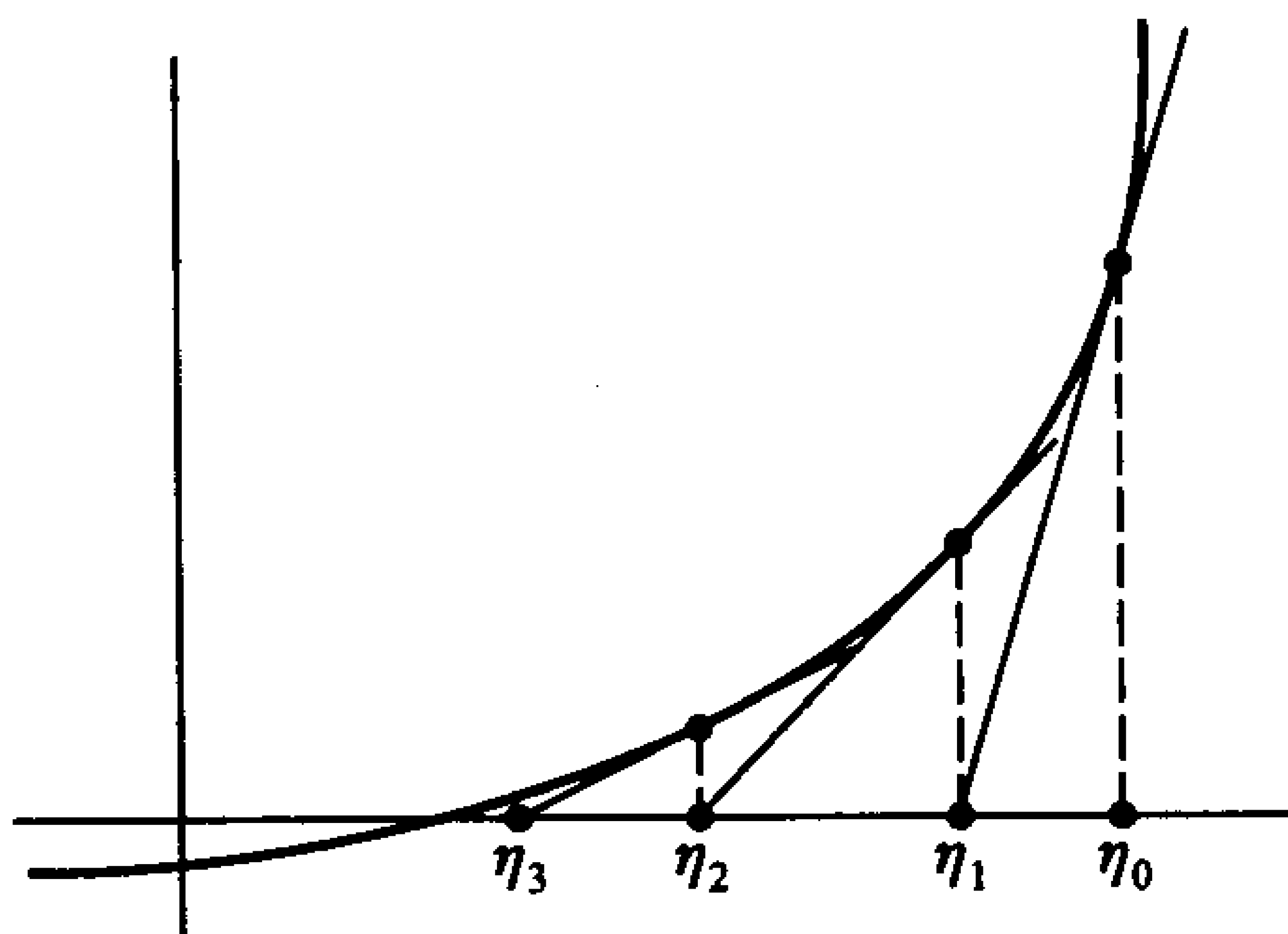


图 4.5

让我们抑制住 ξ 使它在论证中固定不动, 并考虑定义在巴拿赫空间原点的某个邻域中的映射 G. 假设 G 有两个连续的微分. 为明确起见, 我们假设 G 被定义在单位球 B 上, 并假定对每个 $x \in B$, 映射 dG_x 是可逆的, 实际上有

$$\|dG_x^{-1}\| \leqslant K, \quad \|d^2G_x\| \leqslant K.$$

设 $x_0 = 0$, 并假定 x_n 已被定义, 我们令

$$x_{n+1} = x_n - S_n^{-1}G(x_n),$$

其中 $S_n = dG_{x_n}$. 我们要证明, 如果 $\|G(0)\|$ 充分小 (相对 K 而言), 则此过程有明确定义 (即 $\|x_{n+1}\| < 1$) 并快速收敛. 事实上, 如果 τ 为 1 和 2 之间的任意实数 (例如 $\tau = 3/2$), 我们要证明对某个 c(如果 $\|G(0)\|$ 是小的它可以是大的) 有

$$\|x_n - x_{n-1}\| \leqslant e^{-c\tau^n}. \qquad (*)$$

注意, 如果我们对充分大的 c 能建立 $(*)$, 则得到了 $\|x_n\| \leqslant 1$. 事实上,

$$\begin{aligned}\|x_j\| &\leqslant \sum_1^j e^{-c\tau^n} \leqslant \sum_1^\infty e^{-c\tau^n} \leqslant \sum_1^\infty e^{-cn(\tau-1)} \\ &= \frac{e^{-c(\tau-1)}}{1-e^{-c(\tau-1)}},\end{aligned}$$

如果 c 较大, 它则小于 1. 让我们用归纳法来试图证明 $(*)$. 假定它对 n 为真, 我们由泰勒定理有

$$\begin{aligned}\|x_{n+1} - x_n\| &= \|S_n^{-1}G(x_n)\| \\ &\leqslant K\|G(x_{n-1}) - S_{n-1}^{-1}G(x_{n-1})\| \\ &\leqslant K\{\|G(x_{n-1}) - dG_{x_{n-1}}S_{n-1}^{-1}G(x_{n-1})\| \\ &\quad + K\|x_n - x_{n-1}\|^2\}.\end{aligned}$$

现在此不等式右端的第一项为 0, 从而有

$$\|x_{n+1} - x_n\| \leqslant K^2\|x_n - x_{n-1}\|^2 \leqslant K^2 e^{-2c\tau^n}.$$

要使归纳法可以进行下去我们必须要有

$$K^2 e^{-2c\tau^n} \leqslant e^{-c\tau^{n+1}}$$

或者

$$K^2 \leqslant e^{c(2-\tau)\tau^n}. \qquad (**)$$

由于 $\tau < 2$, 最后面的这个不等式可选 c 充分大而得到. 我们还必须验证 $(*)$ 对 $n = 1$ 成立. 这就是说

$$\|S_0^{-1}G(0)\| \leqslant e^{-c\tau}$$

或者

$$\|G(0)\| \leqslant \frac{e^{-c\tau}}{K}. \qquad (***)$$

概括地说, 对 $1<\tau<2$, 选取 c 使得 $K^2 \leqslant e^{(2-\tau)c\tau}$ 和

$$\frac{e^{-c(\tau-1)}}{1-e^{-c(\tau-1)}} \leqslant 1.$$

于是, 如果 $(***)$ 成立, 序列 $\{x_n\}$ 则指数式收敛, 即 $(*)$ 成立. 如果 $x=\lim x_i$, 则 $G(x)=\lim G(x_n)=\lim S_n(x_{n+1}-x_n)=0$, 这就是牛顿法.

作为 c 和 τ 的一种可能的选取, 让 $\tau=\frac{3}{2}$, c 由 $K^2=e^{3c/4}$ 给出, 故 $(**)$ 刚好成立. 我们也可假定 $K>2^{3/4}$, 使得 $e^{3c/4} \geqslant 4^{3/4}$ 或 $e^c \geqslant 4$, 这保证了 $e^{-c/2} \leqslant \frac{1}{2}$, 从而表明 $e^{-c/2}/(1-e^{-c/2}) \leqslant 1$. 于是 $(***)$ 成为要求 $G(0) \leqslant K^{-5}$.

我们以一种最简单情形下的不动点迭代过程的例子来结束本节, 这就是逆映射定理的例子. 假定 $H(0)=0$, dH_0^{-1} 存在, 我们想要在零的附近逆转 H, 即对方程 $H(\eta)-\xi=0$ 用 ξ 解出 η. 上面的理论告诉我们对应于 ξ 的 η 是压缩 $K(\xi,\eta)=\eta-T^{-1}H(\eta)+T^{-1}(\xi)$ 的不动点, 其中 $T=dH_0$. 为使我们的例子尽可能简单, 我们将取 H 为从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$, 并使 $dH_0=I$. 又为了避免指标, 我们将使用混杂的记号 $\boldsymbol{x}=\prec x,y \succ$, $\boldsymbol{u}=\prec u,v \succ$.

考虑映射 $\boldsymbol{x}=H(\boldsymbol{u})$, 其定义为 $x=u+v^2$, $y=u^3+v$. 雅可比矩阵

$$\begin{bmatrix} 1 & 2v \\ 3u^2 & 1 \end{bmatrix}$$

显然在原点为单位矩阵. 又, 在表达式 $K(\boldsymbol{x},\boldsymbol{u})=\boldsymbol{x}+\boldsymbol{u}-H(\boldsymbol{u})$ 中, 差 $H(\boldsymbol{u})-\boldsymbol{u}$ 正好是函数 $J(\boldsymbol{u})=\prec v^2,u^3 \succ$. 一阶项的消去是下面事实的实际表现, 即在形成 $K(\xi,\eta)=\eta-T^{-1}G(\xi,\eta)$ 中, 我们已经使在"中心点"(这里是原点) 有 $dK^2=0$ 了. 我们自然地从 $\boldsymbol{u}_0=0$ 开始做迭代, 而我们的不动点序列为

$$\boldsymbol{u}_1=K(\boldsymbol{x},\boldsymbol{u}_0)=K(\boldsymbol{x},0),\cdots,\boldsymbol{u}_n=K(\boldsymbol{x},\boldsymbol{u}_{n-1}).$$

因而 $\boldsymbol{u}_0=0, \boldsymbol{u}_n=K(\boldsymbol{x},\boldsymbol{u}_{n-1})=\boldsymbol{x}-J(\boldsymbol{u}_{n-1})$, 给出

$$\begin{aligned}
u_1 &= x, & v_1 &= y,\\
u_2 &= x-y^2, & v_2 &= y-x^3,\\
u_3 &= x-(y-x^3)^2 & v_3 &= y-(x-y^2)^3,\\
u_4 &= x-[y-(x-y^2)^3]^2, & v_4 &= y-[x-(y-x^3)^2]^3.
\end{aligned}$$

我们被保证说, 此序列 $\boldsymbol{u}_n$ 在起始点 $\boldsymbol{x}$ 充分靠近 0 时几何地收敛, 而且似乎清楚地表明这两个多项式序列正是在计算反函数 $u(x,y)$ 和 $v(x,y)$ 的泰勒级数展开

式. 我们将要求读者在一个习题中证明此点. 这两个泰勒级数开始于

$$u(x,y) = x - y^2 - 2yx^3 + \cdots$$

$$v(x,y) = y - x^3 + 3x^2y^2 + \cdots .$$

习题

9.1 设 B 为一个赋范线性空间的紧子集, 使得 $rB \subset B$ 对所有 $r \in [0,1]$ 成立. 假设 $F: B \to B$ 是一个常数为 1 的利普希茨映射 (即 $\|F(\xi) - F(\eta)\| \leqslant \|\xi - \eta\|$ 对所有 $\xi, \eta \in B$ 成立.) 证明 F 有一个不动点.[提示: 首先考虑 $G = rF, 0 < r < 1$.]

9.2 给出例子以证明在上面习题中的不动点可以不是惟一的.

9.3 设 X 为一紧度量空间, 并设 $K: X \to X$"削减每个非零距离". 就是说, $\rho(K(x), K(y)) < \rho(x,y), x \neq y$. 证明 K 有一个惟一的不动点.(证明, 不然的话, $\text{glb}\{\rho(K(x), x)\}$ 为正值, 并作为极小值被达到, 然后得到矛盾.)

9.4 设 K 为从 $S \times X$ 到 X 的映射, 其中 X 为完备度量空间而 S 为任一度量空间, 并假定 $K(s,x)$ 是一致于 x 上的对 s 的压缩映射, 并对 x 是 x 的一致利普希茨连续的. 证明不动点 p_s 是 s 的一个利普希茨连续函数. [提示: 修改一下定理 9.1 的推论 4 的证明中的 ε, δ - 起始语句.]

9.5 设 D 为巴拿赫空间 V 中的开子集, $K: D \to V$ 使得 $I - K$ 为具常数 1/2 的利普希茨映射.

(a) 证明如果 $B_r(\alpha) \subset D$ 及 $\beta = K(\alpha)$, 则 $B_{r/2}(\beta) \subset K[D]$.(将不动点定理的一个推论用于某个简单的压缩映射.)

(b) 得出结论说 K 为单射并有开的值域, 以及 K^{-1} 为具有常数 2 的利普希茨映射.

9.6 由前面习题的结果推导出第三章习题 3.20 结果的一个改进形式.

9.7 在定理 9.3 的条件下, 证明如果 $\|dK^2_{\prec\mu,\nu\succ}\| < 1$ 则 $dG^2_{\prec\mu,\nu\succ}$ 可逆.(不要被记号搞糊涂了. 我们仅仅想要知道如果 $\|I - T^{-1} \circ S\| < 1$ 则 S 是可逆的.)

9.8 在第三章定理 11.2 和本章定理 9.3 的陈述之间有一个微小的差异. 在一种情形中我们断言了从一个球 M 出发的一个惟一的连续映射的存在性, 而在另一种场合则是从球 M 出发到球 N 的. 证明这两个连续的解必在 M 上一致从而表明可以去掉值域要在 N 中的要求.(利用定理 9.3 的点对点的惟一性.)

9.9 由在定理 9.4 中的恒等式 $G(\xi, F(\xi)) = 0$ 计算出 dF_α 的表达式, 并证明如果 K 是连续可微的, 则在解的表达式中涉及到的所有映射都是连续的, 从而 $\alpha \mapsto dF_\alpha$ 也因此而连续.

9.10 回到第 9 节末尾所做出的例子, 用归纳法证明多项式 $u_n - u_{n-1}$ 和 $v_n - v_{n-1}$ 不包含次数小于 n 的项.

9.11 继续上面的习题, 因此而证明由取 u_n 中最多次数为 n 的项定义的幂级数在 0 的一个球中收敛, 而且它的和是 H 的逆映射的第一个分量 $u(x,y)$.

9.12 上面的结论在一般情形也成立. 设 $J = \prec K, L \succ$ 为从 $\mathbb{R}^2$ 中 0 的一个球到 $\mathbb{R}^2$ 的任一映射, 它由收敛幂级数

$$K(x,y) = \sum a_{ij}x^i y^j, \quad L(x,y) = \sum b_{ij}x^i y^j$$

定义, 其中不含 0 次和 1 次的项. 约定 $\boldsymbol{x} = \prec x, y \succ$ 和 $\boldsymbol{u} = \prec u, v \succ$, 考虑迭代序列

$$\boldsymbol{u}_0 = 0, \quad \boldsymbol{u}_n = \boldsymbol{x} - J(\boldsymbol{u}_{n-1}).$$

对于当一个幂级数被代入另一个时所出现的情形作出必要的假定, 并用归纳法证明 $\boldsymbol{u}_n - \boldsymbol{u}_{n-1}$ 不包含次数小于 n 的项, 因而 $\boldsymbol{u}_n$ 定义了一个收敛幂级数, 其和为在 0 的一个邻域中逆于 H 的函数 $\boldsymbol{u}(x,y) = \prec u(x,y), v(x,y) \succ$.[记住 $J(\eta) = H(\eta) - \eta$.]

9.13 设 A 为一个巴拿赫代数并设 x 是 A 中一个范数小于 1 的元. 证明

$$(e-x)^{-1} = \prod_{i=1}^{\infty}(1 + x^{2^i}).$$

这意味着如果 π_n 是部分积 $\prod_1^n(1+x^{2^i})$, 则 $\pi_n \to (e-x)^{-1}$.[提示: 用归纳法证明 $(e-x)\pi_{n-1} = e - x^{2^n}$.]

像正文中的牛顿法一样, 这是按指数速度收敛的另一个例子.

4.10 参数弧的积分

本节中我们将做完备性的最后一个应用. 我们首先证明一个非常广泛的扩张定理, 而后应用它到黎曼积分的构造上, 这时的积分是作为对阶梯函数所定义的一个初等积分的扩张.

定理 10.1 *设 U 是赋范线性空间 V 的一个子空间, T 为从 U 到一个巴拿赫空间 W 的一个有界线性映射. 于是 T 有一个到有界线性变换 S 的惟一确定的扩张, 其中 S 是从闭包 $\overline{U}$ 到 W 的映射, 另外, $\|S\| = \|T\|$.*

证明 固定 $\alpha \in \overline{U}$ 并选取 $\{\xi_n\} \subset U$ 使 $\xi_n \to \alpha$. 于是 $\{\xi_n\}$ 是柯西的, 从而 $\{T(\xi_n)\}$ 也是柯西的 (由第 7 节的引理), 故而 $\{T(\xi_n)\}$ 收敛于某个 $\beta \in W$.

如果 $\{\eta_n\}$ 是 U 中另一个收敛于 α 的序列, 则 $\xi_n - \eta_n \to 0$, $T(\xi_n) - T(\eta_n) = T(\xi_n - \eta_n) \to 0$, 故也有 $T(\eta_n) \to \beta$. 因此 β 与序列的选取无关, 并且可以清楚地看到, β 必对 T 的任何一个连续扩张 S 等于 S 在 α 的值 $S(\alpha)$. 如果 $\alpha \in U$, 则 $\beta = \lim T(\alpha_n) = T(\alpha)$, 这是由 T 的连续性得到的. 因此 S 在要求为 T 的连续扩张的条件下在 $\overline{U}$ 上被惟一确定.

剩下来要证明 S 为线性且以 $\|T\|$ 为界. 对任意 $\alpha, \beta \in \overline{U}$, 我们选取 $\{\xi_n\}$, $\{\eta_n\} \subset U$, 使得 $\xi_n \to \alpha, \eta_n \to \beta$. 于是 $x\xi_n + y\eta_n \to x\alpha + y\beta$, 故而

$$
\begin{aligned}
S(x\alpha + y\beta) &= \lim T(x\xi_n + y\eta_n) = x\lim T(\xi_n) + y\lim T(\eta_n) \\
&= xS(\alpha) + yS(\beta).
\end{aligned}
$$

因此 S 为线性. 最后

$$
\|S(\alpha)\| = \lim\|T(\xi_n)\| \leqslant \|T\|\lim\|\xi_n\| = \|T\|\ \|\alpha\|.
$$

因此 $\|T\|$ 为 S 的一个界, 但因为 S 包括了 T, 故 $\|S\| = \|T\|$. □

上面的定理有许多应用, 但我们只用它一次以得到黎曼积分 $\int_a^b f(t)dt$, 其中连续函数 f 映射一个闭区间 $[a,b]$ 到一个巴拿赫空间 W, 而此积分被当作阶梯函数的平凡积分的扩张. 如果 W 是个赋范线性空间, $f: [a,b] \to W$ 是个定义在闭区间 $[a,b] \subset \mathbb{R}$ 上的一个连续函数, 我们可以期待着能够定义 $\int_a^b f(t)dt$ 为 W 中一个适当的向量, 并能够写单实变向量函数的积分学. 直到现在我们还没有这样做, 因为我们需要 W 的完备性来证明这种积分存在!

首先我们只积分某种被称作阶梯函数的初等函数. $[a,b]$ 的一个有限子集 A, 它包含了两个端点 a 和 b, 则称 A 为 $[a,b]$ 的一个划分. 因此 A 是某个有限序列 (的值域)$\{t_i\}_0^n$, 其中 $a = t_0 < t_1 < \cdots < t_n = b$, 从而 A 剖分 $[a,b]$ 为更小区间的序列. 明确地说, 我们取开区间 $(t_{i-1}, t_i), i = 1, \cdots, n$, 为此剖分的区间. 如果 A 和 B 都是剖分且 $A \subset B$, 我们则说 B 是 A 的一个*细分*. B - 剖分的每个区间 (s_{j-1}, s_j) 被包含在 A - 剖分的某个区间 (t_{i-1}, t_i) 中; t_{i-1} 是 A 中小于或等于 s_{j-1} 的最大元, 而 t_i 是大于或等于 s_j 的最小元. 一个*阶梯函数*只不过是个映射 $f: [a,b] \to W$, 它在某个剖分 $A = \{t_i\}_0^n$ 的区间上为*常值*. 那就是说, 存在一个向量序列 $\{\alpha_i\}_1^n$ 使得当 $\xi \in (t_{i-1}, t_i)$ 时 $f(\xi) = \alpha_i$. f 在剖分点的值可以是这些值中的一个也可以不同.

在每个阶梯函数 f, 我们定义 $\int_a^b f(t)dt$ 为 $\sum_{i=1}^n \alpha_i \Delta t_i$, 其中在 (t_{i-1}, t_i) 上 $f = \alpha_i$, 而 $\Delta t_i = t_i - t_{i-1}$. 如果 f 为实函数, 这只简单地是 f 的图像与 t 轴之间

区域做成的矩形面积的和. 现在 f 可用许多不同剖分被描述为阶梯函数. 例如, 如果 f 在 A 的区间上为常值, 而且如果 B 是由 A 添加一个新点 s 得到, 则 f 在 B 的 (更小的) 区间为常值. 我们必须要确信 f 的积分值在改变所描述的剖分时不变. 在刚刚提到的这种情形容易明白. 这一个新点 s 位于某个区间 (t_{i-1}, t_i) 中, 这个区间是由剖分 A 定义的. 这个区间对 A - 和的贡献为 $\alpha_i(t_i - t_{i-1})$, 而在 B - 和中它分裂为 $\alpha_i(t_i - s) + \alpha_i(s - t_{i-1})$. 然而这是同一向量. 而其余的和因子在两个和中是一样的, 从而此积分不变. 一般地, 假设 f 是关于 A 的一个阶梯函数, 同时也是关于 C 的阶梯函数. 那么令 $B = A \cup C$, 它是 A 和 C 的 "共同细分". 我们可以由一系列的步骤从 A 过渡到 B, 其中每一步我们添加一个新的点. 正如我们已经看到的, 在这些步骤中的每一步, 此积分都不变, 故它对 A 和 B 都是一样的. 对 C 和 B 的情形是相似地也一样, 从而对 A 和 C 也一样. 我们已经证明了 $\int_a^b f$ 与用来定义 f 的剖分无关.

现在固定 $[a, b]$ 和 W, 并设 ε 为由 $[a, b]$ 到 W 的所有阶梯函数的集合. 那么, ε 是个向量空间, 因为, 如果 f 和 $g \in \varepsilon$ 为相对于剖分 A 和 B 的阶梯函数, 于是它们在 $C = A \cup B$ 的区间上都是常数, 因此 $xf + yg$ 也是如此. 又, 如果 $C = \{t_i\}_0^n$, 并且如果在 (t_{i-1}, t_i) 上我们有 $f = \alpha_i, g = \beta_i$, 故 $xf + yg = x\alpha_i + y\beta_i$, 因此方程

$$\sum_{i=1}^{n}(x\alpha_i + y\beta_i)\Delta t_i = x\left(\sum_{1}^{n}\alpha_i\Delta t_i\right) + y\left(\sum_{1}^{n}\beta_i\Delta t_i\right)$$

恰好是 $\int_a^b (xf + yg) = x\int_a^b f + y\int_a^b g$. 映射 $f \mapsto \int_a^b f$ 因而从 ε 到 W 是线性的. 最后,

$$\left\|\int_a^b f\right\| = \left\|\sum_{1}^{n}\alpha_i\Delta t_i\right\| \leqslant \sum_{1}^{n}\|\alpha_i\|\Delta t_i \leqslant \|f\|_\infty(b-a),$$

其中 $\|f\|_\infty = \mathrm{lub}\{\|f(t)\| : t \in [a, b]\} = \max\{\|\alpha_i\| : 1 \leqslant i \leqslant n\}$, 就是说我们在 ε 上使用由 W 的范数所定义的一致范数, 于是线性映射 $f \mapsto \int_a^b f$ 以 $(b-a)$ 有界. 如果 W 为完备, 则此变换由定理 10.1 有一个到 $\overline{\varepsilon}$ 上的惟一有界线性扩张, 其中 $\overline{\varepsilon}$ 是 ε 在 $\mathfrak{B}([a, b], W)$ 中取的闭包. 但是我们可以证明 $\overline{\varepsilon}$ 包含了由 $[a, b]$ 到 W 的所有连续函数的空间 $\mathfrak{C}([a, b], W)$ 从而连续函数的积分被惟一地定义.

引理 10.1 $\mathfrak{C}([a, b], W) \subset \overline{\varepsilon}$.

证明 在 $[a, b]$ 上的一个连续函数是一致连续的 (定理 5.1). 就是说, 给出 $\varepsilon^{>0}$, 存在 $\delta^{>0}$, 使得 $|s - t| < \delta \Rightarrow \|f(s) - f(t)\| < \varepsilon$. 现取 $[a, b]$ 的任何一个剖分 $A = \{t_i\}_0^n$, 使得 $\Delta t_i = t_i - t_{i-1} < \delta$ 对所有 i 成立, 并取 α_i 为 f 在 (t_{i-1}, t_i) 中

任一点的值. 则在 $[t_{i-1}, t_i]$ 上 $\|f(t) - \alpha_i\| < \varepsilon$. 因此, 如果 g 是在 $(t_{i-1}, t_i]$ 上取值 α_i 的阶梯函数, 且 $g(a) = \alpha_1$, 则 $\|f - g\|_\infty \leqslant \varepsilon$, 故 f 在 $\overline{\varepsilon}$ 中正是所求. □

我们的主定理是个重新的概述.

定理 10.2 如果 W 是巴拿赫空间, $V = \mathfrak{C}([a,b], W)$ 具一致范数, 则存在一个 $J \in \mathrm{Hom}(V, W)$ 由令 $J(f) = \lim \int_a^b f_n$ 惟一决定, 其中 $\{f_n\}$ 是 ε 中任一收敛于 f 的序列而 $\int_a^b f_n$ 是上面所定义的在 ε 上的积分. 另外又有 $\|J\| \leqslant (b-a)$.

如果 f 从 $[a,b]$ 到 W 是初级的, $c \in [a,b]$, 则 f 在 $[a,c]$ 和 $[c,b]$ 中每个当然也是初级的. 如果 c 被加到剖分 A 中, 而 A 是用来定义 f 的, 并且如果关于 $B = A \cup \{c\}$ 定义 $\int_a^b f$ 的和在 c 被分裂为两个和, 我们显然有 $\int_a^b f = \int_a^c f + \int_c^b f$. 这个相同的等式对于任意 $[a,b]$ 上的连续函数 f 也可得到, 这是因为 $\int_a^b f = \lim \int_a^b f_n = \lim \left(\int_a^c f_n + \int_c^b f_n\right) = \lim \int_a^c f_n + \lim \int_c^b f_n = \int_a^c f + \int_c^b f$.

微积分基本定理依然与我们同在.

定理 10.3 如果 $f \in \mathfrak{C}([a,b], W)$ 且 $F : [a,b] \to W$ 由 $\int_a^x f(t)dt$ 定义, 则 F' 在 (a,b) 中存在且 $F'(x) = f(x)$.

证明 由 f 在 x_0 的连续性, 对每个 ε 存在 δ 使得只要 $|x - x_0| < \delta$, $\|f(x_0) - f(x)\| < \varepsilon$.
从而

$$\left\|\int_{x_0}^{x} (f(x_0) - f(t))dt\right\| \leqslant \varepsilon |x - x_0|,$$

并因 $\int_{x_0}^x f(x_0)dt = f(x_0)(x - x_0)$ 由初级函数积分的定义可以推出, 我们有

$$\left\|f(x_0) - \left(\int_{x_0}^{x} f(t)dt\right) \Big/ (x - x_0)\right\| < \varepsilon.$$

因为 $\int_{x_0}^x f(t)dt = F(x) - F(x_0)$, 这完全等于说 F 的差商收敛于 $f(x_0)$, 这正是我们要证明的. □

习题

10.1 证明下面的定理 10.1 的类比: 设 A 为度量空间 B 的子集, C 为完备度量空间, $F : A \to C$ 一致连续. 则 F 惟一地扩张到从 $\overline{A}$ 到 C 的一个连续映射.

10.2 在习题 7.16 到 7.18 中我们构造了 S 的一个完备化, 即 $\overline{\theta[S]} \subset V^*$. 证明这个完备化在等距同构范围内是惟一的. 就是说, 假设 φ 是 S 是其他的等距嵌入于一个完备空间 X, 证明由 $\varphi \circ \theta^{-1}$ 给出的 S 的两个像的等同映射 (从 $\theta[S]$ 到 $\varphi[S]$) 扩张到从 $\overline{\theta[S]}$ 到 $\overline{\varphi[S]}$ 的一个等距双射.[提示: 应用前面的习题.]

10.3 现在假定 S 是个赋范线性空间 X, 而 X 是完备度量空间 Y 的稠子集. 回想这意味着 Y 中每个点都是子集 X 中一个序列的极限. 证明 X 的向量空间结构可以以惟一的方式扩张使得 Y 成为一个巴拿赫空间. 由于我们从习题 7.18 知道一个度量空间可以完备化, 这再次表明一个赋范线性空间总可以被完备为一个巴拿赫空间.

10.4 在初等微积分中, 如果 f 连续, 则对某个 $xt(a,b)$ 有

$$\int_a^b f(t)dt = f(x)(b-a).$$

证明这对于连续的向量函数 f 不真, 这里的 f 为由

$$f(t) = \prec \sin t, \cos t \succ$$

定义的弧 $f : [0,\pi] \to \mathbb{R}^2$.

10.5 证明取积分和施行线性变换可交换. 就是说, 证明如果 f 是从 $[a,b]$ 到巴拿赫空间 W 的一个连续函数, 又如果 $T \in \mathrm{Hom}(W,X)$, 其中 X 为巴拿赫空间, 则

$$\int_a^b T(f(t))dt = T\left[\int_a^b f(t)dt\right].$$

[提示: 直接从阶梯函数算起.]

10.6 叙述并证明由下面等式所提出的定理:

$$\int_a^b \prec f(t), g(t) \succ dt = \prec \int_a^b f(t)dt, \int_a^b g(t)dt \succ .$$

[提示: 用前一个习题.]

10.7 设 W 为任意赋范线性空间, $\{\alpha_i\}_1^n$ 为 W 中的向量的有限集合, 而 $\{f_i\}_1^n$ 是 $[a,b]$ 上对应的实连续函数的集合. 定义弧 γ 为

$$\gamma(t) = \sum_1^n f_i(t)\alpha_i.$$

证明 $\int_a^b \gamma(t)dt$ 存在并等于

$$\sum_1^n \left[\int_a^b f_i(t)dt\right] \alpha_i.$$

10.8 设 f 为从 $\mathbb{R}^2$ 到巴拿赫空间 W 的一个连续函数. 描述如何可以建立一个二重积分

$$\iint\limits_{I\times J} f(s,t)dsdt$$

的理论, 其中 $I\times J$ 为一个闭的矩形.

10.9 证明如果 f_n 一致收敛于 f 则

$$\int_a^b f_n(t)dt \to \int_a^b f(t)dt.$$

如果你理解了积分的定义和性质, 这则是微不足道的事.

10.10 假设 $\{f_n\}$ 是从 $[a,b]$ 到一巴拿赫空间 W 的光滑弧的序列使得 $\sum_1^\infty f_n'(t)$ 一致收敛. 假设 $\sum_1^\infty f_n(a)$ 也收敛. 证明 $\sum f_n(t)$ 因此而一致收敛且 $f=\sum_1^\infty f_n$ 光滑, 并且 $f'=\sum_1^\infty f_n'$.(用上面的习题和微积分基本定理.)

10.11 证明即使 W 不是个巴拿赫空间, 如果弧 $f:[a,b]\to W$ 有连续的导数, 则 $\int_a^b f'$ 存在并等于 $f(b)-f(a)$.

10.12 设 X 为赋范线性空间, 令 $\prec l,\xi\succ = l(\xi)$, 其中 $\xi\in X$, $l\in X^*$. 现设 f 和 g 分别为由闭区间 $[a,b]$ 到 X 和 X^* 的连续可微函数 (弧). 证明分部积分公式:

$$\prec g(b),f(b)\succ - \prec g(a),f(a)\succ = \int_a^b \prec f(t),g'(t)\succ dt + \int_a^b \prec f'(t),g(t)\succ dt$$

[提示: 应用第三章的定理 8.4.]

10.13 叙述上面分部积分公式可以推广到对任意有界双线性映射 $\omega: V\times W\to X$ 成立, 其中 X 为一个巴拿赫空间.

10.14 设 $t\mapsto l_t$ 为从闭区间 $[a,b]$ 到对偶 W^* 的一个固定的连续映射, W 为巴拿赫空间. 假设对任一从 $[a,b]$ 到 W 的连续函数 g, 我们有

$$\int_a^b g(t)dt=0 \implies \int_a^b l_t(g(t))dt=0.$$

证明存在一个固定的 $L\in W^*$ 使得对所有连续弧 $g:[a,b]\to W$ 有

$$\int_a^b l_t(g(t))dt = L\left(\int_a^b g(t)dt\right).$$

证明由此得到对所有 t, $l_t=L$.

10.15 用上面习题推出 3.15 节的一般欧拉方程.

4.11 复数系

复数系 $\mathbb{C}$ 是继有理数和实数之后必须要研究的第三个基本数域, 读者必定在以往对它已有所接触.

几乎每一个人都把复数ξ 看作是一对实数, 即 ξ 的 "实部和虚部", 因而实数系 $\mathbb{C}$ 被看作是笛卡儿二维空间 $\mathbb{R}^2$, 但具有某些深一层的结构. 特别地, 一个复函数简单地说就是某种向量函数, 等价于一对有序的实函数, 仍为其实部和虚部.

把复数系从它的向量属性 $\mathbb{R}^2$ 区分出来的是由于一个附加的运算即复乘法的出现. $\mathbb{R}^2$ 的向量运算连同它的复乘法运算使 $\mathbb{C}$ 成了一个变换代数. 最终还表示 $\prec 1,0 \succ$ 是 $\mathbb{C}$ 中惟一的乘法单位元, 而且每个非零复数都有一个乘法逆元. 这些添加进来的事实概括起来说就是:$\mathbb{C}$ 是一个域, 它让我们可将 $\mathbb{C}$ 用作向量空间理论中的一个新的纯量域. 事实上, 第一和第二章的全部阐述在把 $\mathbb{R}$ 换做 $\mathbb{C}$ 时仍然有效. 纯量乘法现在是乘以复数. 因此 $\mathbb{C}^n$ 是有序的复数的 n 联组的向量空间, 而一个 n 联组与一个复纯量 α 的积定义为 $\alpha \prec \xi_1,\cdots,\xi_n \succ=\prec \alpha\xi_1,\cdots,\alpha\xi_n \succ$, 其中 $\alpha\xi_i$ 是复数乘法.

到了认真处理复乘法的时候了. 像读者大体知道的那样, 它由一个样子古怪的公式给出, 而这个公式是由把一个元 $\xi=\prec x_1,x_2 \succ$ 想成是 x_1+ix_2 的形式, 其中 $i^2=-1$ 再使用通常的代数法则而引起来的. 于是我们有

$$\begin{aligned}\xi\eta &= (x_1+ix_2)(y_1+iy_2)\\ &= x_1y_1+ix_1y_2+ix_2y_1+i^2x_2y_2\\ &= (x_1y_1-x_2y_2)+i(x_1y_2+x_2y_1),\end{aligned}$$

因此我们的定义是

$$\prec x_1,x_2 \succ\prec y_1,y_2 \succ=\prec x_1y_1-x_2y_2,x_1y_2+x_2y_1 \succ.$$

当然, 必须验证这个运算是交换的并满足一个代数的规则. 直接检验是可以的但显得呆板, 我们将在习题中以一种简洁的方式表明.

映射 $x \mapsto \prec x,0 \succ$ 为域 $\mathbb{R}$ 到域 $\mathbb{C}$ 的一个同构单射. 显然它保持和, 读者可通过默验知道它也保持积. 按约定把 x 等同于它的像 $\prec x,0 \succ$, 故可将 $\mathbb{R}$ 看作 $\mathbb{C}$ 的一个子域.

这个神秘的 i 等同于 $\mathbb{C}$ 中的偶对 $\prec 0,1 \succ$, 由此 $i^2=\prec 0,1 \succ\prec 0,1 \succ=\prec -1,0 \succ$, 后者我们已将其等同于 -1 了. 按这些等同关系我们有 $\prec x,y \succ=$

$\prec x, 0 \succ + \prec 0, y \succ = \prec x, 0 \succ + \prec 0, 1 \succ \prec y, 0 \succ = x + iy$, 这就是从现在起我们要写出复数的方式.

映射 $x + iy \mapsto x - iy$ 是 $\mathbb{C}$ 与自己的一个域同构. 就是说, 它保持了和与积, 这点很容易被验证. 这种自身的同构被称为一个自同构. 前面这个自同构被称为复共轭, $\zeta = x + iy$ 的像 $x - iy$ 被称作 ζ 的共轭, 以 $\overline{\zeta}$ 表示. 我们在习题中将要求读者证明共轭是 $\mathbb{C}$ 的惟一的保持子域 $\mathbb{R}$ 中每个元素固定不变的自同构 (除去恒同自同构).

称 $\zeta = x + iy = \prec x, y \succ$ 的欧几里得范数为 ζ 的绝对值, 以 $|\zeta|$ 表示, 故 $|\zeta| = |x + iy| = (x^2 + y^2)^{1/2}$. 因为最后证明 $|\zeta\gamma| = |\zeta|\,|\gamma|$, 故这样称呼是有道理的. 这个等式可以由取平方和相乘加以验证, 但是下面的做法要简洁得多, 这就是首先注意到绝对值和共轭自同构之间的关系:

$$\zeta\overline{\zeta} = |\zeta|^2$$

$[(x + iy)(x - iy) = x^2 - (iy)^2 = x^2 + y^2]$. 然后, $|\zeta\gamma|^2 = (\zeta\gamma) \cdot (\overline{\zeta\gamma}) = (\zeta\overline{\zeta})(\gamma\overline{\gamma}) = |\zeta|^2|\gamma|^2$, 从而取平方根便给出了我们的等式. 等式 $\zeta\overline{\zeta} = |\zeta|^2$ 也向我们表明了如果 $\zeta \neq 0$, 则 $\overline{\zeta}/|\zeta|^2$ 是其乘法逆元.

因为实数系 $\mathbb{R}$ 是复数系 $\mathbb{C}$ 的一个子域, 故任何 $\mathbb{C}$ 上的向量空间也自动地为 $\mathbb{R}$ 上的向量空间: 乘以复纯量包括了乘以实纯量. 而且任何一个在复向量空间之间的复线性变换自动地是实线性的. 当然反过来并不成立. 例如, 一个从 $\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的实线性映射 T 一般来说不是从 $\mathbb{C}$ 到 $\mathbb{C}$ 的复线性, 一个 $\mathrm{Hom}\mathbb{R}^4$ 的实线性 S 在把 $\mathbb{R}^4$ 看作 $\mathbb{C}^2$ 时也不一定成为 $\mathrm{Hom}\mathbb{C}^2$ 中的复线性映射. 我们在习题中要讨论这个问题.

在复向量空间之间的一个映射 F 的复可微性有显然的定义:$\Delta F_\alpha = T + \mathfrak{o}$, 其中 T 为复线性, 当然由前面的评注知其也是实可微的. 但 F 可能是实可微却不是复可微的. 由在第 8 节末尾的讨论知道, 如果 $\{a_n\} \subset \mathbb{C}$, 并且 $\{|a_n|\delta^n\}$ 有界, 则级数 $\sum a_n\zeta^n$ 在 (实) 巴拿赫代数 $\mathbb{C}$ 的球 $B_\delta(0)$ 上收敛, 而且 $F(\zeta) = \sum_0^\infty a_n\zeta^n$ 在此球上为实可微, 有 $dF_\beta(\zeta) = (\sum_1^\infty na_n\beta^n)\zeta = F'(\beta) \cdot \zeta$. 但乘以 $F'(\beta)$ 显然是个在一维复向量空间 $\mathbb{C}$ 上的复线性运算. 因此, 由收敛的复幂级数定义的复函数是自动复可微的. 然而我们甚至可以走得更远. 这时, 如果 $\zeta \neq 0$, 我们在定义方程

$$\Delta F_\beta(\zeta) = F'(\beta) \cdot \zeta + \mathfrak{o}(\zeta)$$

可以除以 ζ 得到当 $\zeta \to 0$ 时

$$\frac{\Delta F_\beta(\zeta)}{\zeta} \to F'(\beta)$$

的结果. 那就是说, $F'(\beta)$ 现在又是一个纯正的导数了, 它以一个复无穷小 ζ 在差商的分母上.

复可微性的结果是多得不可胜数的, 我们将其绝大部分留到复变函数课程中去体验未来的愉悦. 但是看看第十二章末尾关于留数计算的问题以及在第十一章习题 4.3 中对下面代数基本定理的证明吧.

定理 11.1 *复系数的每个多项式是线性因子的乘积.*

一个较弱但等价的陈述是说, 每个多项式至少有一个 (复) 根. 事情的关键是 x^2+1 在 $\mathbb{R}$ 上不能分解因子 (即没有实根) 而在 $\mathbb{C}$ 上我们有 $x^2+1=(x+i)(x-i)$, 有两个根 $\pm i$.

为以后要用我们多说几句关于复指数函数的话: $\exp\zeta = e^\zeta = \sum_0^\infty \zeta^n/n!$. 如果 $\zeta = x+iy$, 我们有 $e^\zeta = e^{x+iy} = e^x e^{iy}$, 并且 $e^{iy} = \sum_0^\infty (iy)^n/n! = (1-y^2/2!+y^4/4!-\cdots)+i(y-y^3/3!+y^5/5!-\cdots) = \cos y + i\sin y$. 因此 $e^{x+iy} = e^x(\cos y + i\sin y)$. 那就是说, 复函数 $\exp(x+iy)$ 的实部和虚部分别为 $e^x\cos y$ 和 $e^x\sin y$.

习题

11.1 直接从定义证明复乘法的可结合性.

11.2 证明复数的分配律:$\alpha(\xi+\eta)=\alpha\xi+\alpha\eta$.

11.3 证明用一个实数 a 作的纯量乘法 $a \prec x, y, \succ = \prec ax, ay \succ$ 在 $\mathbb{C}=\mathbb{R}^2$ 中与将 a 解释为复数 $\prec a, 0 \succ$ 的复乘法定义是一致的.

11.4 设 θ 为复域的保持实数不动的自同构. 证明 θ 或是恒同映射或是复共轭.[提示: $(\theta(i))^2 = \theta(i^2) = \theta(-1) = -1$. 证明其平方为 -1 的惟一复数 $x+iy$ 为 $\pm i$, 于是得证.]

11.5 如果我们记住 $\mathbb{C}$ 特别是个二维的实向量空间 $\mathbb{R}^2$, 我们则看出用复数 $a+ib$ 去乘 $\mathbb{C}$ 中元必定定义了 $\mathbb{R}^2$ 上的一个线性变换. 证明它的矩阵是

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

11.6 上面的习题提示了复数系可以是所有形如

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

2×2 实矩阵的集合 A. 证明 A 是矩阵代数 $\mathbb{R}^{\bar{2}\times\bar{2}}$ 的子代数 (即 A 在乘法, 加法和纯量乘法下封闭), 并且 $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mapsto a+ib$ 是从 A 到 $\mathbb{C}$ 的一个双射, 它保持了所有的代数运算. 因此我们可以得出结论说一个代数应满足的规则对 $\mathbb{C}$ 自动成立. 为什么?

11.7 在上面复数系的矩阵模式中证明绝对值等式 $|\zeta\gamma| = |\zeta|\,|\gamma|$ 是一个行列式性质.

11.8 设 W 为一实向量空间, V 为实向量空间 $W\times W$. 证明存在 Hom$V$ 中一个 θ 使得 $\theta^2 = -I$.(把 $\mathbb{C}$ 想成为在乘以 i 下的实向量空间 $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R}$)

11.9 设 V 为实向量空间, 并设 $\theta\in \mathrm{Hom}V$ 满足 $\theta^2=-I$. 证明如果 $i\alpha$ 被定义为 $\theta(\alpha)$, 则 V 成为一个复向量空间. 如果复向量空间 V 是由实向量空间 W 按现在这种方式和上面习题那样的方式做成, 我们则称 V 为 W 的复化.我们将把 W 自身看成是 V 的一个子空间 (实际为 $W\times\{0\}$), 并且 $V = W\bigoplus iW$.

11.10 证明复向量空间 $\mathbb{C}^n$ 是 $\mathbb{R}^n$ 的复化. 证明更一般地对任何集 A, 复向量空间 $\mathbb{C}^A$ 是实向量空间 $\mathbb{R}^A$ 的复化.

11.11 设 V 为实向量空间 W 的复化, 定义在 V 上的复共轭运算. 那就是说, 证明存在一个实线性映射 φ 使得 $\varphi^2=I$ 且 $\varphi(i\alpha)=-i\varphi(\alpha)$. 证明, 反之, 如果 V 是个复向量空间, 且 φ 是 V 上的一个共轭 [一个实线性映射 φ 使得 $\varphi^2=I$ 而 $\varphi(i\alpha)=-i\varphi(\alpha)$], 则 V 是 (同构于) 一个实线性空间 W 的复化.(应用第一章的定理 5.5 于等式 $\varphi^2-I=0$.)

11.12 设 W 为实向量空间, V 为其复化. 证明 Hom$W$ 中每个 T"扩张" 到 Hom$V$ 中的复线性映射 S 使它与共轭 φ 交换. 我们所说的 S 扩到了 T 的意思自然是 $S\restriction(W\times\{0\}) = T$. 证明, 反过来, 如果 S 在 Hom$V$ 中与共轭互换, 则 S 是 Hom$W$ 中一个 T 的扩张.

11.13 在这种情形中我们自然把 S 叫做 T 的复化. 证明如果 S 是 T 的复化, 则它的在 V 中的零空间 X 为直和 $X = N\bigoplus iN$, 其中 N 是 T 在 W 中的零空间. 记住我们是把 V 看作 $W\bigoplus iW$ 的.

11.14 在一个复赋范线性空间 V 上, 其范数被要求为复齐次的: 对所有复数 λ 有

$$\|\lambda\alpha\| = |\lambda|\,\|\alpha\|.$$

证明在 $\mathbb{C}^n$ 上的 $\|\ \|_1, \|\ \|_2, \|\ \|_\infty$ 的自然定义都具有这个性质.

11.15 如果一个实赋范线性空间 W 被复化为 $V = W+iW$, 没有显而易见的公式能把 W 上的实范数转换为 V 上的复范数. 但是证明, V(实际上是 $W\times W$) 上任一个乘积范数可以被用来产生一个等价的复范数.[提示: 对给出的 $\prec \xi,\eta \succ\in V$, 考虑数的集合 $\{\|(x+iy)\prec \xi,\eta \succ\|:|x+iy|=1\}$, 并力图从此集合中得到一个可用的一个数.]

11.16 证明每个非零复数有一个对数. 那就是说, 证明如果 $u+iv\neq 0$, 则存在一个 $x+iy$ 使得 $e^{x+iy}=u+iv$. (写出方程 $e^x(\cos y+i\sin y)=u+iv$, 并用点小智慧解出它.)

11.17 代数基本定理和第一章的定理 5.5 表明如果 V 是个复向量空间, 而 $T\in \mathrm{Hom}V$ 满足 $p(T)=0$, 其中 p 是个多项式, 则存在 V 的子空间 $\{V_i\}_1^n$, 复数 $\{\lambda_i\}_1^n$ 和整数 $\{m_i\}_1^n$ 使得 $V=\bigoplus_1^n V_i$, 其中 V_i 对每个 i 是 T - 不变空间, 且对每个 i, 在 V_i 上 $(T-\lambda_i T)^{m_i}=0$. 证明上面所说的确如此. 也证明如果 V 为有限维, 则 $\mathrm{Hom}V$ 中的每个 T 必定满足某个多项式方程 $p(t)=0$. (考虑向量 $I,T,T^2,\cdots,T^{n^2},\cdots$ 在向量空间 $\mathrm{Hom}(V)$ 的线性无关性或线性相关性.)

11.18 假设上面习题中的多项式 p 具有实系数. 利用复共轭是 $\mathbb{C}$ 的自同构这个事实证明如果 λ 是 p 的一个根则 $\bar{\lambda}$ 也是个根.

证明如果 V 是实空间 W 的复化, T 是 $R\in \mathrm{Hom}W$ 的复化, 则存在一个实多项式 p 使得 $p(T)=0$.

11.19 证明如果 W 为有限维实向量空间, $R\in \mathrm{Hom}W$ 为一个同构, 则存在一个 $A\in \mathrm{Hom}W$ 使得 $R=e^A$(即 $\log R$ 存在). 这是个难啃的习题, 但它可通过习题 8.19 到 8.23, 11.12, 11.17 和 11.18 得到证明.

*4.12 弱方法

我们的在有限维空间上所有范数都等价的定理使人想到这种空间上的极限理论应该独立于范数而被接受, 而且我们早先关于具有限维定义域的每个线性变换都自动为有界的定理更加强了这种印象. 我们在本节要研究这个问题. 在某种意义上说, 这样的努力并不合适, 因为我们不能完全没有范数, 还因为它们用起来是如此顺手, 即使我们不必用它时也会使用.

粗略地说, 我们要做的就是用研究整个实 (值) 映射 $\{l\circ F: l\in V^*\}$ 的办法去研究向量 (值) 映射 F.

定理 12.1 如果 V 为有限维, 则在 V 中 $\xi_n\to\xi$(关于任一个因而也是每一个范数) 当且仅当对每个 $l\in V^*$ 在 $\mathbb{R}$ 中 $l(\xi_n)\to l(\xi)$.

证明 如果 $\xi_n \to \xi$ 且 $l \in V^*$, 则因为 l 是自动连续的有 $l(\xi_n) \to l(\xi)$. 反之, 如果对 V^* 中每个 l 有 $l(\xi_n) \to l(\xi)$, 则当在 V 中选取一组基 $\{\beta_i\}_1^n$ 时, 我们在对偶基中每个泛函 ε_i 有 $\varepsilon_i(\xi_n) \to \varepsilon_i(\xi)$, 而这表明在相关的 1 范数下 $\xi_n \to \xi$, 这是因为有 $\|\xi_n - \xi\|_1 = \sum_1^n |\varepsilon_i(\xi_n) - \varepsilon_i(\xi)| \to 0$. □

注 如果 V 为任意赋范线性空间, 故 $V^* = \mathrm{Hom}(V, \mathbb{R})$ 表示了有界线性泛函的集合, 那么我们说 ξ_n弱趋向于 ξ 是指对每个 $l \in V^*, l(\xi_n) \to l(\xi)$. 上面的定理重新组词叙述便是: *在一个有限维空间中弱收敛性和范数收敛性是等价的概念.*

我们将看到, 以一种相似的方式即使用 V^* 中的泛函和 V^{**} 与 V 的自然同构, 可以把参数弧的积分和微分通通回溯到实变量的实函数的标准微积中. 因而, 如果 $f \in \mathfrak{C}([a,b], V)$, $\lambda \in V^*$, 则 $\lambda \circ f \in \mathfrak{C}([a,b], \mathbb{R})$, 故由标准微积分知道 $\int_a^b \lambda \circ f$ 存在. 如果让 λ 变化, 则可验证出映射 $\lambda \mapsto \int_a^b \lambda \circ f$ 是线性的, 从而在 V^{**} 中, 因此由 V 中一个惟一确定的向量 α 给出 (由对偶性; 见第二章定理 3.2). 那就是说, 存在惟一的 $\alpha \in V$ 使得 $\lambda(\alpha) = \int_a^b \lambda \circ f$ 对所有 $\lambda \in V^*$ 成立, 我们便定义这个 α 为 $\int_a^b f$. 因此积分被定义成使得它与线性泛函的作用可交换. $\int_a^b f$ 是那样的向量, 使得对所有 $\lambda \in V^*$,

$$\lambda\left(\int_a^b f\right) = \int_a^b \lambda(f(t))dt.$$

相似地, 如果所有的实函数 $\{\lambda \circ f : \lambda \in V^*\}$ 都在 x_0 可微, 则映射 $\lambda \mapsto (\lambda \circ f)'(x_0)$ 由标准微积分中导数的线性性知也是线性的:

$$\begin{aligned}((c_1\lambda_1 + c_2\lambda_2) \circ f)' &= (c_1(\lambda_1 \circ f) + c_2(\lambda_2 \circ f))' \\ &= c_1(\lambda_1 \circ f)' + c_2(\lambda_2 \circ f)'.\end{aligned}$$

因此, 再次存在惟一 $\alpha \in V$ 使得对所有 $\lambda \in V^*$,

$$(\lambda \circ f)'(x_0) = \lambda(\alpha),$$

而且如果我们定义这个 α 为导数 $f'(x_0)$ 的话, 我们再一次定义了一个微积分的运算, 它与线性泛函具有可换性:

$$(\lambda \circ f')(x_0) = (\lambda \circ f)'(x_0).$$

现在微积分的基本定理表现如下.

如果 $F(x)=\int_a^x f$, 则 $(\lambda\circ F)(x)=\int_a^x \lambda\circ f$, 这是积分的弱定义. 标准微积分的基本定理于是说 $(\lambda\circ F)'$ 存在且 $(\lambda\circ F)'(x)=(\lambda\circ f)(x)=\lambda(f(x))$. 由导数的弱定义我们便有 F' 存在且 $F'(x)=f(x)$.

我们不能用弱方法轻易得到的一个结果是范数不等式 $\|\int_a^b f\|\leqslant(b-a)\|f\|_\infty$. 这需要一个关于有限维空间上范数的一个定理, 在本教程中不打算证明它.

定理 12.2 $\|\alpha^{**}\|=\|\alpha\|$ 对所有 $\alpha\in V$ 成立.

我们要断言的是 $\text{lub}\{|\alpha^{**}(\lambda)|/\|\lambda\|\}=\|\alpha\|$. 因为 $\alpha^{**}(\lambda)=\lambda(\alpha)$, 又因为 $|\lambda(\alpha)|\leqslant\|\lambda\|\cdot\|\alpha\|$(由 $\|\lambda\|$ 的定义), 故有

$$\text{lub}|\alpha^{**}(\lambda)|/\|\lambda\|\leqslant\|\alpha\|.$$

因而我们的问题便是找到一个 $\lambda\in V^*$ 使得 $\|\lambda\|=1$ 且 $|\lambda(\alpha)|=\|\alpha\|$. 如果乘以一个适当的常数 (以 $c\alpha$ 代替 $\alpha, c=1/\|\alpha\|$), 我们可假设 $\|\alpha\|=1$. 于是 α 在一个单位球面上, 问题便是找一个泛函 $\lambda\in V^*$ 使得仿射子空间 $\lambda=1$(超平面) 与单位球在 α 相切 (故 $\lambda(\alpha)=1$), 否则将在单位球外 (故当 $\|\xi\|=1$ 时 $|\lambda(\xi)|\leqslant 1$, 从而 $\|\lambda\|\leqslant 1$). 从几何上讲显然这样的 "切平面" 必定存在, 但我们就此而止了.

如果我们承认这个定理, 则由于

$$\begin{aligned}\left|\lambda\left(\int_a^b f\right)\right|&=\left|\int_a^b \lambda(f(t))dt\right|\leqslant(b-a)\max\{|\lambda(f(t))|:t\in[a,b]\}\\&\leqslant(b-a)\|\lambda\|\max\{\|f(t)\|\}\quad(\text{由}|\lambda(\alpha)|\leqslant\|\lambda\|\cdot\|\alpha\|)\\&=(b-a)\|\lambda\|\cdot\|f\|_\infty,\end{aligned}$$

我们得到

$$\begin{aligned}\left\|\int_a^b f\right\|=\left\|\left(\int_a^b f\right)^{**}\right\|&=\text{lub}\left|\lambda\left(\int_a^b f\right)\right|/\|\lambda\|\\&\leqslant(b-a)\|f\|_\infty,\end{aligned}$$

它两端的项构成了所要的不等式.

第五章　内积空间

在这短短的一章里我们将研究在 2 范数下发生了些什么, 而我们将发现这里打开了线性分析的一个全新的分支. 这些范数可被抽象地描述为具有从纯量积产生的特性. 它们是普通几何长度在有限和无限维情形的类比, 它们实际上都具有欧氏几何的所有概念, 像两个向量之间夹角的概念, 垂直性 (正交性) 以及毕达哥拉斯定理, 以及多刚体运动的存在性.

这个额外结构的影响对无穷维空间情形特别具有戏剧性. 无限正交基在极大范围内都存在, 而且能像在有限维空间中基那样容易处理, 虽然这时一个向量的基展开式是一个收敛无穷级数 $\xi = \sum_1^\infty x_i\alpha_i$. 数学中许多最重要的级数展式都是这些正交基展开的例子. 举例说, 在下一章我们将看到在 $[0,\pi]$ 上一个连续函数 f 的傅里叶级数展式是 f 在 2 - 范数$\|f\|_2 = \left(\int_0^\pi f^2\right)^{1/2}$ 下对一组特定的正交基 $\{\alpha_n\}_1^\infty = \{\sin nt\}_1^\infty$ 的基展开式. 如果一个空间在纯量积范数下是完备的, 则称它为一个*希尔伯特空间*.这样的空间的更高等的理论是数学的最漂亮的部分之一.

5.1　内积 (纯量积)

在一个实向量空间 V 上的一个*内积*是从 $V \times V$ 到 $\mathbb{R}$ 的一个实函数, 它在偶对 $\prec \xi,\eta \succ$ 上的值通常用 (ξ,η) 表示, 使得

(a) 当 η 固定时 (ξ,η) 对 ξ 是线性的;

(b) $(\xi,\eta)=(\eta,\xi)$(对称性);

(c) 如果 $\xi\neq 0$, 则 $(\xi,\xi)>0$(正定性).

如果将 (c) 换成较弱的条件:

(c′) 对所有 $\xi\in V$ 有 $(\xi,\xi)\geqslant 0$,

则称 (ξ,η) 为半内积.

内积的两个重要例子是

$$(\boldsymbol{x},\boldsymbol{y})=\sum_1^n x_iy_i,\quad \text{其中}V=\mathbb{R}^n$$

和

$$(f,g)=\int_a^b f(t)g(t)dt,\quad \text{其中}V=\mathfrak{C}([a,b]).$$

在复向量空间上, 条件 (b) 必须被换为

(b′) $(\xi,\eta)=\overline{(\eta,\xi)}$(埃尔米特对称),

这里的横线 '—' 代表复共轭. 相应的例子是 $(\boldsymbol{z},\boldsymbol{w})=\sum_1^n z_i\overline{w}_i$, 这里的 $V=\mathbb{C}^n$, 以及当 V 为 $[a,b]$ 上的连续复函数时, $(f,g)=\int_a^b f\overline{g}$. 我们只讨论实的情形.

由 (a) 和 (b) 得到一个半内积当第一个变量不动时对第二个变量也是线性的, 因此是一个对称的双线性函数, 它的相伴的二次形式 $q(\xi)=(\xi,\xi)$ 是正定或半正定的 [(c) 或 (c′), 见第二章最后一节].

形式 q 的正负定性质有着深远的结果, 我们马上就能看到这点.

定理 1.1 *施瓦兹不等式*

$$|(\xi,\eta)|\leqslant(\xi,\xi)^{1/2}(\eta,\eta)^{1/2}$$

对任何半内积成立.

证明 我们有 $0\leqslant(\xi-t\eta,\xi-t\eta)=(\xi,\xi)-2t(\xi,\eta)+t^2(\eta,\eta)$ 对所有 $t\in\mathbb{R}$ 成立. 因为它对 t 是二次的永不负的, 它不能有不同的根, 从而通常的 (b^2-4ac) 公式表明 $4(\xi,\eta)^2-4(\xi,\xi)(\eta,\eta)\leqslant 0$, 它等价于施瓦兹不等式. □

我们也可直接进行. 如果 $(\eta,\eta)>0$, 并且如果令 $t=(\xi,\eta)/(\eta,\eta)$ 代入证明中第一行的二次不等式中, 则所得结果简化后便是施瓦兹不等式. 如果 $(\eta,\eta)=0$, 则 (ξ,η) 也必为零 (否则最初的不等式显然对某个 t 不成立), 那么施瓦兹不等式不用说是对的.

推论 1.1 如果 (ξ,η) 为内积, 则 $\|\xi\|=(\xi,\xi)^{1/2}$ 是一个范数.

证明

$$\begin{aligned}\|\xi+\eta\|^2&=(\xi+\eta,\xi+\eta)\\&=\|\xi\|^2+2(\xi,\eta)+\|\eta\|^2\leqslant\|\xi\|^2+2\|\xi\|\,\|\eta\|+\|\eta\|^2\end{aligned}$$

(由施瓦兹不等式)

$$=(\|\xi\|+\|\eta\|)^2,$$

这证明了三角不等式. 另外, $\|c\xi\|=(c\xi,c\xi)^{1/2}=(c^2(\xi,\xi))^{1/2}=|c|\,\|\xi\|$. □

注意, 施瓦兹不等式 $|(\xi,\eta)|\leqslant\|\xi\|\,\|\eta\|$ 现在正好是说双线性泛函 (ξ,η) 关于内积范数以 1 为有界.

一个以内积范数为范数的赋范线性空间 V 被称作*准希尔伯特空间*.如果在此范数下 V 为完备, 它则是一个*希尔伯特空间*.

前面提到的两个内积的例子第一次给了我们 2 - 范数真正的解释:

$$\|\boldsymbol{x}\|_2=\left(\sum_1^n x_i^2\right)^{1/2},\boldsymbol{x}\in\mathbb{R}^n$$

和

$$\|f\|_2=\left(\int_a^b f^2\right)^{1/2},f\in\mathfrak{C}([a,b])$$

是内积范数.

因为 $\mathbb{R}^n$ 上的内积范数在与 n 维欧几里得空间在笛卡儿坐标对应下成为欧几里得长度, 习惯上称 $\mathbb{R}^n$ 本身也为欧几里得 n 维空间 $\mathbb{E}^n$, 但这时我们总理解它使用了内积范数.

一个有限维向量空间在任何内积范数下都是一个希尔伯特空间, 这是因为有限维数的性质保证了它的完备性. 另一方面, 我们将在习题 1.10 中看到, $\mathfrak{C}([a,b])$ 在 2 范数下是不完备的, 因而在此范数下只是个准希尔伯特空间而不是希尔伯特空间. (但是回想在一致范数 $\|f\|_\infty$ 下 $\mathfrak{C}([a,b])$ 是完备的.) 真正在数学中使用希尔伯特空间的重要之处在于任意一个准希尔伯特空间都可以被完备化为一个希尔伯特空间, 但无穷维希尔伯特空间理论大部分都超出了本书的范围.

内积范数从某种意义上说具有最可能光滑的单位球面, 因为这些球面是二次曲面.

正交性给了准希尔伯特理论一种特殊的性质. 两个向量 α 和 β 被称作正交的并记为 $\alpha\perp\beta$ 是说如果有 $(\alpha,\beta)=0$. 这个定义从几何中获取了灵感; 在第一章我们曾说过, 两个几何向量为垂直的当且仅当它们的坐标三联组 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 满足 $(\boldsymbol{x},\boldsymbol{y})=0$. 这是个可进一步探索的令人感兴趣的问题; 由余弦定律 $(c^2=a^2+b^2-2ab\cos\theta)$ 表明两个几何向量之间的夹角由 $(\boldsymbol{x},\boldsymbol{y})=\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta$ 给出. 这会启发我们去定义在准希尔伯特空间中两个向量 ξ 和 η 之间的夹角 θ 为 $(\xi,\eta)=\|\xi\|\,\|\eta\|\cos\theta$, 但我们不会用这个更一般的公式化陈述.

我们说两个子集 A 和 B 的正交的并记为 $A\perp B$ 是说, 如果对每个 $\alpha\in A$ 和每个 $\beta\in B$ 有 $\alpha\perp\beta$; 对任意子集 A, 我们令 $A^\perp=\{\beta\in V:\beta\perp A\}$.

引理 1.1 *如果 β 正交于集合 A, 则 β 正交于 $\overline{L(A)}$, 它表示 A 的线性张成空间的闭包. 由此得到对任何子集 B, $B^\perp$ 为闭子空间.*

证明 第一个论断有赖于内积对其的一个变量的线性性和连续性; 留给读者去证明. 至于 $A=B^\perp$, 它包含了它自己线性张成空间的闭包, 这由第一部分推出; 故它是个闭子空间. □

引理 1.2 *在任一个准希尔伯特空间中我们有平行四边形定律*

$$\|\alpha+\beta\|^2+\|\alpha-\beta\|^2=2(\|\alpha\|^2+\|\beta\|^2)$$

以及毕达哥拉斯定理,

$$\alpha\perp\beta\text{当且仅当}\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2.$$

如果 $\{\alpha_i\}_1^n$ 是一个 (两两) 正交的向量组, 则

$$\left\|\sum_1^n\alpha_i\right\|^2=\sum_1^n\|\alpha_i\|^2.$$

证明 由内积 $\|\alpha+\beta\|^2$ 的双线性性 $\|\alpha+\beta\|^2=\|\alpha\|^2+2(\alpha,\beta)+\|\beta\|$, 因而看出 $\|\alpha+\beta\|^2=\|\alpha\|^2+\|\beta\|^2$ 当且仅当 $(\alpha,\beta)=0$, 它就是毕达哥拉斯定理. 写出类似的 $\|\alpha-\beta\|^2$ 的展开式并与上面展开式相加, 我们便得到了平行四边形定律. 最后一个陈述归纳地来自毕达哥拉斯定理和引理 1.1. 或者我们可用展开左端的内积并留意到所有 "混合项" 均由正交性消去而直接得到. □

读者会留意到在此引理中没有用到施瓦兹不等式，但是要在证明 $\|\xi\| = (\xi,\xi)^{1/2}$ 是个范数之前就叙述此引理就显得不合情理了.

如果 $\{\alpha_i\}_1^n$ 正交且非零，那么 $\|\sum_1^n x_i\alpha_i\|^2 = \sum_1^n x_i^2\|\alpha_i\|^2$ 表明只有所有的系数 x_i 为零时 $\sum_1^n x_i\alpha_i$ 才能为零. 因此,

推论 1.2 一个 (两两) 正交非零向量的有限组是无关的. 相似地, 正交子空间的有限组也是无关的.

习题

1.1 完成定理 1.1 的第二个证明.

1.2 重新考查定理 1.1 的证明并证明如果 ξ 和 η 无关, 则施瓦兹不等式是严格的.

1.3 继续上面的习题, 现在证明如果 ξ 和 η 无关则三角不等式是严格的.

1.4 (a) 证明两个半内积的和是个半内积.

(b) 证明如果 (μ,ν) 是向量空间 W 上的半内积, 并且如果 T 是从向量空间 V 到 W 的线性变换, 则 $[\xi,\eta]=(T\mu,T\nu)$ 是 V 上的一个半内积.

(c) 从 (a) 和 (b) 推导出

$$(f,g)=f(a)g(a)+\int_a^b f'(t)g'(t)dt$$

是 $V=\mathfrak{C}^1([a,b])$ 上的一个半内积. 证明它是个内积.

1.5 如果 α 被固定, 我们知道 $f(\xi)=(\xi,\alpha)$ 为连续. 为什么? 证明一般地作为 $V\times V$ 到 $\mathbb{R}$ 的映射 (ξ,η) 为连续.

1.6 设 V 为一个二维希尔伯特空间, 并设 $\{\alpha_1,\alpha_2\}$ 为 V 的任一组基. 证明内积 (ξ,η) 有形式

$$(\xi,\eta)=ax_1y_1+b(x_1y_2+x_2y_1)+cx_2y_2,$$

其中 $b^2<ac$. 当然, 这里有 $\xi=x_1\alpha_1+x_2\alpha_2, \eta=y_1\alpha_1+y_2\alpha_2$.

1.7 证明如果 $\omega(\boldsymbol{x},\boldsymbol{y})=ax_1y_1+b(x_1y_2+x_2y_1)+cx_2y_2$ 及 $b^2<ac$, 则 ω 是 $\mathbb{R}^2$ 上的一个内积.

1.8 设 $\omega(\xi,\eta)$ 是在一个有限维向量空间 V 上的任意对称的双线性泛函, 并设 $q(\xi)=\omega(\xi,\xi)$ 为其相伴的二次型. 证明对 V 的任一组选取的基, 方程 $q(\xi)=1$ 成为 ξ 坐标 $\{x_i\}$ 的二次方程.

1.9 详细证明如果向量 β 在准希尔伯特空间中垂直于集合 A, 则 β 正交于 $\overline{L(A)}$.

1.10 从前一章我们知道 [0,1] 上实阶梯函数的一致极限的集合 $\mathcal{E}$ 上可定义黎曼积分, 而且 $\mathcal{E}$ 包含了所有的连续函数. 给出阶梯函数 k, 它在 $[0, \frac{1}{2}]$ 上取值 1 而在 $[\frac{1}{2}, 1]$ 上取值 0, 证明对任意连续函数 f 有 $\|f-k\|_2 > 0$. 但是, 证明存在一个连续函数的序列 $\{f_n\}$ 使得 $\|f_n - k\|_2 \to 0$. 因而由证明上面的序列 $\{f_n\}$ 是柯西的但在 $\mathfrak{C}([0,1])$ 中并不收敛而表明 $\mathfrak{C}([0,1])$ 在 2 - 范数下是不完备的.

5.2 正交投影

在几何推理中最重要的手段之一是从一点向一直线或平面 "作垂线" 然后用直角三角形论证. 在准希尔伯特空间理论中这个手段同等地重要. 如果 M 是个子空间, α 为 V 中任一元, 则 "从 α 向 M 作垂线的垂足" 的意思是指 M 中向量 μ 使得 $(\alpha-\mu)\perp M$(如果这样的 μ 存在; 见图 5.1). 将 α 写成 $\mu+(\alpha-\mu)$, 我们看出对每个 V 中 α 存在 M 中的 "垂足"μ 等价于直和分解 $V = M \bigoplus M^{\perp}$. 现在准确地说, 是希尔伯特空间的完备性保证了这个直和分解, 这我们很快就会看到. 我们由证明一个几何直观事实开始, 即 μ 是从 α 引向 M 垂线的垂足当且仅当 μ 是 M 中最靠近 α 的点.

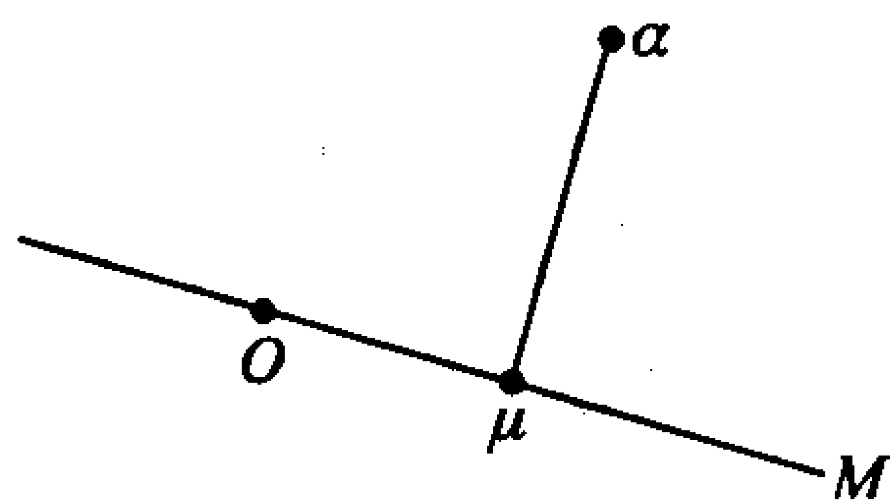


图 5.1

引理 2.1 *如果 μ 在子空间 M 中, 则 $(\alpha-\mu)\perp M$ 当且仅当 μ 是 M 中最靠近 α 的惟一的点, 即 μ 是在 M 中对 α 的 "最佳近似".*

证明 如果 $(\alpha-\mu)\perp M$, 又 ξ 为 M 中任意另外一个点, 则 $\|\alpha-\xi\|^2 = \|(\alpha-\mu)+(\mu-\xi)\|^2 = \|\alpha-\mu\|^2 + \|\mu-\xi\|^2 > \|\alpha-\mu\|^2$. 因此 μ 是 M 中惟一最靠近 α 的点. 反之, 假定 μ 是 M 中最靠近 α 的点, 并设 ξ 为 M 中任一非零向量. 于是 $\|\alpha-\mu\|^2 \leqslant \|(\alpha-\mu)+t\xi\|^2$, 当右端展开后它成为 $0 \leqslant 2t(\alpha-\mu,\xi)+t^2\|\xi\|^2$. 这时所有 t 都成立仅当 $(\alpha-\mu,\xi)=0$(否则让 $t=?$). 因此 $(\alpha-\mu)\perp M$. □

在这个引理的基础上可以清楚看出寻找 μ 的一个办法是取 M 中一个序列 μ_n 使得 $\|\alpha-\mu_n\| \to \rho(\alpha, M)$, 并希望定义 μ 为它的极限. 关键之处是: 我们能证明这样的序列 $\{\mu_n\}$ 总是柯西的, 但是如果 M 不是完备的则此极限可能不存在!

引理 2.2 *如果 $\{\mu_n\}$ 是子空间 M 中一个序列, 它与某个向量 α 的距离收敛于从 α 到 M 的距离 ρ, 则 $\{\mu_n\}$ 是柯西的.*

证明 由平行四边形定律, $\|\mu_n - \mu_m\|^2 = \|(\alpha - \mu_n) - (\alpha - \mu_m)\|^2 = 2(\|\alpha - \mu_n\|^2 + \|\alpha - \mu_m\|^2) - \|2\alpha - (\mu_n + \mu_m)\|^2$. 因为右端第一项当 $m, n \to \infty$ 时收敛于 $4\rho^2$, 而且因为第二项总是 $\leqslant -4\rho^2$(提出因子 2), 我们看出当 $n, m \to \infty$ 时 $\|\mu_n - \mu_m\|^2 \to 0$. □

定理 2.1 如果若 M 是准希尔伯空间 V 的一个完备子空间, 则 $V = M \bigoplus M^{\perp}$. 特别这个结论对准希尔伯特空间的任意有限维子空间和对希尔伯特空间的任一个闭子空间是成立的.

证明 立刻由最后两个引理得到, 因为这时 $\mu = \lim\mu_n$ 存在, $\|\alpha - \mu\| = \rho(\alpha, M)$, 故 $(\alpha - \mu) \perp M$. □

如果 $V = M \bigoplus M^{\perp}$, 则沿 $M^{\perp}$ 到 M 上的投影被称为到 M 上的正交投影, 或者简单地称为到 M 的投影, 这是因为在到 M 的相伴 M 的各种补空间的投影中正交投影是最重要的. 因此, 如果 M 是 V 的完备子空间, 并且如果 P 是到 M 上的投影, 则 $P(\xi)$ 既是从 ξ 向 M 引垂线的垂足 (这是 "投影" 这个词的来源) 又是在 M 中对 ξ 的最好近似 (引理 2.1).

引理 2.3 如果 $\{M_i\}_1^n$ 是完备的两两正交子空间的有限汇集, 并且如果对 V 中向量 α , α_i 是 α 在 M_i 上的投影 $i = 1, \cdots, n$, 则 $\sum_1^n \alpha_i$ 是 α 在 $\bigoplus_1^n M_i$ 上的投影.

证明 我们必须证明 $\alpha - \sum_1^n \alpha_i$ 正交于 $\bigoplus_1^n M_j$, 而这只要证明它分别正交于每个 M_i 即可. 但是如果 $\xi \in M_j$ 则 $(\alpha - \sum_1^n \alpha_i, \xi) = (\alpha - \alpha_j, \xi)$, 其中对 $i \neq j$ 有 $(\alpha_i, \xi) = 0$ 而且因 α_j 是 α 在 M_j 上的投影, 故 $(\alpha - \alpha_j, \xi) = 0$. 因此 $(\alpha - \sum_1^n \alpha_i, \xi) = 0$. □

引理 2.4 ξ 在由单个非零向量 η 张成的一维空间上的投影为 $((\xi, \eta)/\|\eta\|^2)\eta$.

证明 这里的 μ 必定为某个形如 $x\eta$ 的元. 但 $(\xi - x\eta) \perp \eta$ 当且仅当

$$0 = (\xi - x\eta, \eta) = (\xi, \eta) - x\|\eta\|^2 \quad \text{或} \quad x = (\xi, \eta)/\|\eta\|^2.$$ □

我们称数 $(\xi, \eta)/\|\eta\|^2$ 为 ξ 的 η -傅里叶系数.如果 η 是个单位 (或法化) 向量, 则此傅里叶系数正好是 (ξ, η). 由引理 2.3 知, 如果 $\{\varphi_i\}_1^n$ 是一组正交的非零向量, 并且如果 $\{x_i\}_1^n$ 为对应于它们的 ξ 的傅里叶系数, 则 $\sum_1^n x_i\varphi_i$ 是向量 ξ

在由 $\{\varphi_i\}_1^n$ 张成的子空间 M 上的投影. 因此, $\xi - \sum_1^n x_i\varphi_i \perp M$, 而且 (引理 2.1) $\sum_1^n x_i\varphi_i$ 是在 M 中对 ξ 的最佳近似. 如果 ξ 就在 M 中, 则这些论述说的都是 $\xi = \sum_1^n x_i\varphi_i$. (这个当然可以直接验证, 只要设 $\xi = \sum_1^n a_i\varphi_i$ 为 ξ 的对基的展开式, 并计算 $(\xi, \varphi_j) = \sum_1^n a_i(\varphi_i, \varphi_j) = a_j\|\varphi_j\|^2$.)

如果一个向量的正交集合 $\{\varphi_i\}$ 也同时是法化的 ($\|\varphi_i\| = 1$), 我们则称此集为法正交的.

定理 2.2 如果 $\{\varphi_i\}_1^\infty$ 是一无穷法正交序列, 并且如果 $\{x_i\}_1^\infty$ 是向量 ξ 对此序列相应的傅里叶系数, 则

$$\sum_1^\infty x_i^2 \leqslant \|\xi\|^2 \quad (\text{贝塞尔不等式})$$

并且 $\xi = \sum_1^\infty x_i\varphi_i$ 当且仅当 $\sum_1^\infty x_i^2 = \|\xi\|^2$(帕塞瓦尔方程)

证明 令 $\sigma_n = \sum_1^n x_i\varphi_i$, $\xi = (\xi - \sigma_n) + \sigma_n$, 并记住有 $\xi - \sigma_n \perp \sigma_n$, 我们便有

$$\|\xi\|^2 = \|\xi - \sigma_n\|^2 + \sum_1^n x_i^2.$$

因此, $\sum_1^n x_i^2 \leqslant \|\xi\|^2$ 对所有 n 成立, 故证明了贝塞尔不等式, 并且 $\sigma_n \to \xi$(即 $\|\xi - \sigma_n\| \to 0$) 当且仅当 $\sum_1^n x_i^2 \to \|\xi\|^2$, 这证明了帕塞瓦尔等式. □

我们称形式级数 $\sum x_i\varphi_i$ 为 ξ 的**傅里叶级数**(关于法正交集 $\{\varphi_i\}$ 的). 帕塞瓦尔条件说, ξ 的傅里叶级数收敛于 ξ 当且仅当 $\|\xi\|^2 = \sum_1^\infty x_i^2$.

称一个无穷法正交序列 $\{\varphi_i\}_1^\infty$ 为一个准希尔伯特空间 V 的一个基是说, 如果 V 中每个元都是它的傅里叶级数的和.

定理 2.3 一个无穷法正交序列 $\{\varphi_i\}_1^\infty$ 是准希尔伯特空间 V 的一个基当 (且仅当) 它的线性张成空间稠于 V.

证明 如果 ξ 为 V 中任一元, 并设 $\{x_i\}$ 为其傅里叶系数的序列, 由于 $\{\varphi_i\}$ 的线性张成空间稠于 V, 故任给 ε, 存在一个有限线性组合 $\sum_1^n y_i\varphi_i$ 它以 ε 近似于 ξ. 但是由引理 2.3 和 2.1 知, $\sum_1^m x_i\varphi_i$ 是在 $\{\varphi_i\}_1^m$ 的张成空间中对 ξ 的最佳近似, 故对任意 $m \geqslant n$ 有

$$\|\xi - \sum_1^m x_i\varphi_i\| \leqslant \varepsilon,$$

即 $\xi = \sum_1^\infty x_i\varphi_i$. □

推论 2.1 如果 V 是希尔伯特空间, 则法正交序列 $\{\varphi_i\}_1^\infty$ 为一个基当且仅当 $\{\varphi_i\}^\perp = \{0\}$.

证明 设 M 为 $\{\varphi_i\}_1^\infty$ 的线性张成空间的闭包. 因为 $V \doteq M + M^\perp$, 并由引理 1.1 知 $M^\perp = \{\varphi_i\}^\perp$, 我们看出 $\{\varphi_i\}^\perp = \{0\}$ 当且仅当 $V = M$, 那么由定理, 它成立当且仅当 $\{\varphi_i\}$ 是个基. □

注意, 当只用正交基时, 一个向量 ξ 在一个基元 β 的系数总是傅里叶系数 $(\xi, \beta)/\|\beta\|^2$. 因此 ξ 的 β - 系数只依赖于 β 而与 β 之外的基元的选取无关. 但是从第二章我们知道, 当包含 β 的一个任意基被用到时, 则 ξ 的 β 系数随基而变化. 这就部分地解释了正交基得到偏爱的原因.

我们常常由"正交化"某个序列得到一组法正交序列.

引理 2.5 如果 $\{\alpha_i\}$ 是一个有限或无穷的独立向量序列, 则存在一个法正交序列 $\{\varphi_i\}$ 使得 $\{\alpha_i\}_1^n$ 和 $\{\varphi_i\}_1^n$ 对所有 n 都张成相同的线性空间.

证明 由于法化是小事一桩, 我们只做正交化. 为了确定性, 假设此序列是无穷的, 并设 M_n 由 $\{\alpha_1, \cdots, \alpha_n\}$ 线性张成. 设 μ_n 为 α_n 在 M_{n-1} 上的正交投影, 并令 $\varphi_n = \alpha_n - \mu_n$, 其中 $\varphi_1 = \alpha_1$. 这就是我们要的序列. 我们有 $\varphi_i \in M_i \subset M_{n-1}$, 其中 $i < n$, 并且 $\varphi_n \perp M_{n-1}$, 故这些向量 φ_i 相互正交. 另外, 因为 α_n 不在 M_{n-1} 中, 故 $\varphi_n \neq 0$. 于是由引理 1.2 的推论知, $\{\varphi_i\}_1^n$ 是 n 维向量空间 M_n 的一个无关的子集, 故 $\{\varphi_i\}_1^n$ 张成 M_n. □

实际的正交化序列 $\{\varphi_n\}$ 的计算可以递推地进行, 从 $\varphi_1 = \alpha_1$ 开始, 并注意到由于 μ_n 是 α_n 在 $\varphi_1, \cdots, \varphi_{n-1}$ 张成的空间上的投影, 它必为向量 $\sum_1^{n-1} c_i\varphi_i$, 其中 c_i 为 α_n 关于 φ_i 的傅里叶系数.

例如考虑在 $\mathfrak{C}([0,1])$ 中的序列 $\{x^n\}_0^\infty$. 我们有 $\varphi_1 = \alpha_1 = 1$. 其次, $\varphi_2 = \alpha_2 - \mu_2 = x_2 - c \cdot 1$, 其中

$$c = (\alpha_2, \varphi_1)/\|\varphi_1\|^2 = \int_0^1 x \cdot 1 \Big/ \int_0^1 (1)^2 = \frac{1}{2}.$$

然后 $\varphi_3 = \alpha_3 - (c_2\varphi_2 + c_1\varphi_1)$, 其中

$$c_1 = \int_0^1 x^2 \cdot 1 \Big/ \int_0^1 (1)^2 = \frac{1}{3},$$

和

$$c_2 = \int_0^1 x^2\left(x-\frac{1}{2}\right)\bigg/\int_0^1\left(x-\frac{1}{2}\right)^2 = \left(\frac{1}{4}-\frac{1}{6}\right)\bigg/\left(\frac{2}{24}\right) = 1.$$

因此 $\{x^n\}_0^\infty$ 在 $\mathfrak{C}([0,1])$ 中正交化的前三项是 $1, x-\frac{1}{2}$, $x^2-(x-\frac{1}{2})-\frac{1}{3}=x^2-x+\frac{1}{6}$. 这个过程完全是初等的, 但在仅仅算了几项之后这个计算便显然成了沉重的负担了.

我们记得在一般的双线性理论中, 如果对 $\beta\in V$, 我们定义 $\theta_\beta: V\to\mathbb{R}$ 为 $\theta_\beta(\xi)=(\xi,\beta)$, 则 $\theta_\beta\in V^*$ 而 $\theta:\beta\mapsto\theta_\beta$ 便是 V 到 V^* 的一个线性映射. 如果 (ξ,η) 是个内积, 则 $\theta_\beta(\beta)=\|\beta\|^2>0$, 其中 $\beta\neq 0$, 故 θ 是个单射. 实际上由我们要求读者要证明的习题中知 θ 是个等距映射. 如果 V 是有限维的, θ 的单射性表明 θ 是个同构. 然而我们有一个非常惊人的结果:

定理 2.4 *θ 是个同构当且仅当 V 是希尔伯特空间.*

证明 首先假定 V 是个希尔伯特空间. 我们必须证明 θ 是满射, 即每个 V^* 中的非零元 F 都是 θ_β 的形式. 给出这样的 F, 令 N 为其零空间, α 为一与 N 正交的向量 (定理 2.1), 并考虑 $\beta=c\alpha$, c 待定. 每个 V 中向量 ξ 可惟一表成和 $\xi=x\beta+\eta$, $\eta\in N$,[这只是说 V/N 是一维的, 它是我们假定知道了的事实. 但是我们也可直接用 F 去检验. 这时看出 $F(\xi-x\beta)=0$ 当且仅当 $x=F(\xi)/F(\beta)$.] 但是现在方程

$$F(\xi)=F(x\beta+\eta)=xF(\beta)=xcF(\alpha)$$

和

$$\theta_\beta(\xi)=(\xi,\beta)=(x\beta+\eta,\beta)=x\|\beta\|^2=xc^2\|\alpha\|^2$$

表明如果我们取 $c=F(\alpha)/\|\alpha\|^2$, 则 $\theta_\beta=F$.

反之, 如果 θ 是满射的 (并假定为等距的), 则它在 $\mathrm{Hom}(V,V^*)$ 中是个同构, 因为由第四章定理 7.6 知 V^* 是完备的, 从而由同一章的定理 7.3 推出 V 是完备的. 证完. □

习题

2.1 在引理 2.1 的证明中, 如果 $(\alpha-\mu,\xi)\neq 0$, t 的什么值会与不等式 $0\leqslant 2t(\alpha-\mu,\xi)+t^2\|\xi\|^2$ 相矛盾?

2.2 证明定理 2.3 中的 "仅当" 的部分.

2.3 设 $\{M_i\}$ 为准希尔伯特空间 V 中的完备子空间的正交序列, 设 P_i 为在 M_i 上的 (正交) 投影. 证明对 V 中任意 ξ, $\{P_i\xi\}$ 是柯西序列.

2.4 证明函数 $\{\sin nt\}_{n=1}^{\infty}$ 构成了准希尔伯特空间 $\mathfrak{C}([0,\pi])$ 一系列的正交元, 其中为标准内积 $(f,g)=\int_0^{\pi}f(t)g(t)dt$. 又证明 $\|\sin nt\|_2=\sqrt{\pi/2}$.

2.5 对上题中 $\mathfrak{C}([0,\pi])$ 的一组正交集计算函数 $f(t)=t$ 的傅里叶系数. 在由 $\sin t$ 和 $\sin 2t$ 张成的二维空间中什么是对 t 的最佳 2 范数近似? 描画这个近似函数的图像, 以通常微积分曲线作图的方式表现出它的突出特征.

2.6 由在 $[0,\pi/2]$ 上 $f(t)=\pi/2$, 在 $(\pi/2,\pi]$ 上 $f(t)=0$ 定义的 "阶梯" 函数 f 自然在 $\pi/2$ 不连续. 尽管如此, 在 $\mathfrak{C}([0,\pi])$ 中关于 $\{\sin nt\}_1^{\infty}$ 计算出 f 的傅里叶系数, 并画出在 $\{\sin nt\}_1^3$ 张成的空间它的最佳近似的图像.

2.7 证明函数 $\{\sin nt\}_{n=1}^{\infty}\bigcup\{\cos nt\}_{n=0}^{\infty}$ 在准希尔伯特空间 $\mathfrak{C}([-\pi,\pi])$ 中形成元素的正交集, 其中为标准内积 $(f,g)=\int_{-\pi}^{\pi}f(t)g(t)dt$.

2.8 计算 $\{x^n\}_0^{\infty}$ 在 $\mathfrak{C}([-1,1])$ 中的正交化的前三项.

2.9 利用有界线性变换的范数定义和施瓦兹不等式证明 $\|\theta_\beta\|\leqslant\|\beta\|$(其中 $\theta_\beta(\xi)=(\xi,\beta)$). 为了得到 $\beta\mapsto\theta_\beta$ 是等距的结论, 我们也需要相反的不等式: $\|\theta_\beta\|\geqslant\|\beta\|$. 用 ξ 的一个特殊值证明它.

2.10 证明如果 V 是个不完备的准希尔伯特空间, 则 V 有一个真闭子空间 M, 使得 $M^{\perp}=\{0\}$.[*提示: 必存在 $F\in V^*$ 而不具形式 $F(\xi)=(\xi,\alpha)$.*] 连同定理 2.1, 这表明了一个准希尔伯特空间 V 为希尔伯特空间当且仅当 $V=M\bigoplus M^{\perp}$ 对所有闭子空间成立.

2.11 等距映射 $\theta:\alpha\mapsto\theta_\alpha$[其中 $\theta_\alpha(\xi)=(\xi,\alpha)$] 将准希尔伯特空间嵌入到它的共轭空间 V^*. 我们已经知道 V^* 是完备的. 为什么? 作为 V^* 子空间的 V 的闭包因而也是完备的, 因而我们可以将 V 完备为一个巴拿赫空间. 设 H 为其完备化. 它包含了 V(它的等距像) 作为稠子空间. 证明 V 上的内积惟一地扩张到 H 从而 H 上的范数为这扩张了的内积范数时, H 是个希尔伯特空间.

2.12 证明准希尔伯特空间 V 在 $\alpha\mapsto\theta_\alpha$ 的等距嵌入到 V^* 下的正交性等价于 2.3 节所讨论的零化. 讨论在零化子 A° 的性质和本章引理 1.1 之间的关联.

2.13 证明如果 C 是准希尔伯特空间 V 的一个非空完备的凸子集, 并且如果 α 是不在 C 中的任一向量, 则存在一个惟一的 $\mu\in C$ 最靠近 α.(检查引理 2.2 的证明.)

5.3 自伴变换

定义 3.1 *如果 V 是个准希尔伯特空间, 那么, 称 $\mathrm{Hom}V$ 中的 T 是自伴的是说对每个 $\alpha,\beta\in V$ 如果有 $(T\alpha,\beta)=(\alpha,T\beta)$. 所有自伴变换的集合以 SA 表示.*

自伴性使我们想到 T 应该在单射 $\theta: V \to V^*$ 下成为它自己的伴随. 我们现在来检验一下. 因为 $(\alpha, \beta) = \theta_\beta(\alpha)$, 我们可以重写方程 $(T\alpha, \beta) = (\alpha, T\beta)$ 为 $\theta_\beta(T\alpha) = \theta_{T\beta}(\alpha)$, 又由 T^* 的定义, $(T^*(\theta_\beta))(\alpha) = \theta_\beta(T\alpha)$. 这对所有 α, β 成立当且仅当 $T^*(\theta_\beta) = \theta_{T\beta}$, 或者说 $T^* \circ \theta = \theta \circ T$, 这就是我们所断言的等同性.

引理 3.1 如果 V 为有限维希尔伯特空间, 且 $\{\varphi_i\}_1^n$ 是 V 的一个法正交基, 则 $T \in \mathrm{Hom}V$ 为自伴当且仅当 T 关于 $\{\varphi_i\}$ 的矩阵 $\{t_{ij}\}$ 是对称的 $(\boldsymbol{t} = \boldsymbol{t}^*)$

证明 如果我们代入 α 和 β 的基展开式并展开, 我们得到对所有 α 和 β 有 $(\alpha, T\beta) = (T\alpha, \beta)$ 成立当且仅当 $(\varphi_i, T\varphi_j) = (T\varphi_i, \varphi_j)$ 对所有 i 和 j 成立. 但是 $T\varphi_l = \sum_{k=1}^n t_{kl}\varphi_k$, 将其代入前面最后一个等式的内积中, 此方程成为 $t_{ij} = t_{ji}$. 即, T 为自伴当且仅当 $\boldsymbol{t} = \boldsymbol{t}^*$. □

一个自伴的 T 被称作非负的是指如果 $(T\xi, \xi) \geqslant 0$ 对所有 ξ 成立. 于是 $[\xi, \eta] = (T\xi, \eta)$ 是个半内积!

引理 3.2 如果 T 是个非负自伴变换, 则对所有 ξ 有 $\|T(\xi)\| \leqslant \|T\|^{1/2} (T\xi, \xi)^{1/2}$. 因此如果 $(T\xi, \xi) = 0$ 则 $T(\xi) = 0$, 从而更一般地, 如果 $(T\xi_n, \xi_n) \to 0$ 则 $T(\xi_n) \to 0$.

证明 如果 T 同时为非负和自伴, 则 $[\xi, \eta] = (T\xi, \eta)$ 是个半内积, 故由施瓦兹不等式,

$$|(T\xi, \eta)| = |[\xi, \eta]| \leqslant [\xi, \xi]^{1/2}[\eta, \eta]^{1/2} = (T\xi, \xi)^{1/2}(T\eta, \eta)^{1/2}.$$

取 $\eta = T\xi$, 右边的因子成为 $(T(T(\xi)), T\xi)^{1/2}$, 那么由施瓦兹不等式和 $\|T\|$ 的定义知它小于或等于 $\|T\|^{1/2}\|T(\xi)\|$. 除以 $\|T\xi\|$ 便得引理中的不等式. □

若 $\alpha \neq 0$ 且 $T(\alpha) = c\alpha$, 对某个 c 成立, 则称 α 是 T 的特征向量(eigenvector, proper vector, characteristic vector), 称 c 为相伴的特征值 (eigenvalue, proper value, characteristic value).

定理 3.1 如果 V 是有限维希尔伯特空间, T 是 $\mathrm{Hom}V$ 中的一个自伴元, 则 V 有一个完全由 T 的特征向量组成的法正交基.

证明 考虑函数 $(T\xi, \xi)$. 它是 ξ 的连续实函数, 并在单位球 $S = \{\xi : \|\xi\| = 1\}$ 上被 $\|T\|$ 界于上 (由施瓦兹不等式). 令 $m = \mathrm{lub}\{(T\xi, \xi) : \|\xi\| = 1\}$. 因为 S 为

紧 (有界闭集), $(T\xi,\xi)$ 在 S 上某点 α 具有值 m. 现在, $m-T$ 是个非负自伴的变换 (验证!), 而且 $(T\alpha,\alpha)=m$ 等价于 $((m-T)\alpha,\alpha)=0$. 因此由引理 3.2 有 $(m-T)\alpha=0$, 从而 $T\alpha=m\alpha$. 于是我们找到了 T 的一个特征向量. 现在令 $V_1=V,\alpha_1=\alpha,m_1=m$, 并设 V_2 为 $\{\alpha_1\}^{\perp}$. 于是 $T[V_2]\subset V_2$, 因为如果 $\xi\perp\alpha_1$, 则 $(T\xi,\alpha_1)=(\xi,T\alpha_1)=m(\xi,\alpha_1)=0$.

因此我们可以对 T 在希尔伯特空间 V_2 上的限制重复上面的论证, 从而找到 V_2 中的 α_2 使得 $\|\alpha_2\|=1$ 和 $T(\alpha_2)=m_2\alpha_2$, 其中 $m_2=\mathrm{lub}\{(T\xi,\xi):\|\xi\|=1$ 且 $\xi\in V_2\}$. 显然 $m_2\leqslant m_1$. 然后我们令 $V_3=\{\alpha_1,\alpha_2\}^{\perp}$, 继续做下去, 最后得到由 T 的特征向量 $\{\alpha_i\}_1^n$ 组成的法正交基. □

现在设 $\lambda_1,\cdots,\lambda_r$ 为在一连串数 $m_1,\cdots,m_n$ 中不同的值, 并设 M_j 为那些 $m_i=\lambda_j$ 对应的基向量 α_i 线性张成的空间. 于是这些子空间 M_j 相互正交, $V=\bigoplus_1^r M_j$, 每个 M_j 为 $T-$ 不变, 而 T 在 M_j 上的限制是 λ_i 乘上单位变换. 因为 M_j 中所有的非零向量都是以 λ_i 为特征值的特征向量, 所以如果这些张成 M_j 的 α_i 被换成 M_j 的任何其他法正交基, 我们仍旧有一个由特征向量组成的法正交基. 这些 α_i 因此一般来说并不是惟一决定的. *但是子空间 M_j 和特征值 λ_j 是惟一的.* 如果我们证明每个特征向量必在一个 M_j 中则就可以得出这个结论.

引理 3.3 *在上面讨论的背景中, 如果 $\xi\neq 0$ 且 $T(\xi)=x\xi$ 对某个 $x\in\mathbb{R}$ 成立, 则 $\xi\in M_j$(因而 $x=\lambda_j$) 对某个 j 成立.*

证明 因为 $V=\bigoplus_1^r M_j$, 我们有 $\xi=\sum_1^r\xi_i$, $\xi_i\in M_i$. 于是 $\sum_1^r x\xi_i=x\xi=T(\xi)=\sum_1^r T(\xi_i)=\sum_1^r\lambda_i\xi_i$ 从而 $\sum_1^r(x-\lambda_i)\xi_i=0$. 因为子空间 M_i 独立的, 故每个分量 $(x-\lambda_i)\xi_i$ 为 0. 但是由于 $\xi\neq 0$ 故某个 $\xi_j\neq 0$. 因此, $x=\lambda_j$, 而 $i\neq j$ 时 $\xi_i=0$, 故 $\xi=\xi_j\in M_j$. □

我们已经证明了下面的定理.

定理 3.2 *如果 V 是有限维希尔伯特空间, T 是 $\mathrm{Hom}V$ 的一个自伴元, 则存在 V 的惟一确定的子空间 $\{V_i\}_1^r$ 和不同的纯量 $\{\lambda_i\}_1^r$ 使得 $\{V_i\}$ 为一正交族, 其和为 V, 而且 T 在 V_i 上的限制是 λ_i 乘单位变换.*

如果 V 是有限维向量空间且给出了 $T\in\mathrm{Hom}V$, 则我们知道如何计算相关联的映射, 诸如 T^2 和 T^{-1}(如果存在的话) 以及向量 $T\alpha,T^{-1}\alpha$ 等等, 这只要对

V 选取一组基, 然后计算矩阵的积, 逆 (当存在时) 等等. 某些这种计算, 特别涉及到逆映射是相当艰难的. 由 T 的特征向量组成的基的一大好处在于它把所有这些计算变得平凡易行.

要明白这点, 设 $\{\beta_n\}$ 为 V 的完全由 T 的特征向量组成的基, 设 $\{r_n\}$ 为对应的特征值. 要计算 $T\xi$, 我们先写出对 ξ 的基展开式, $\xi = \sum_1^n x_i\beta_i$, 于是 $T\xi = \sum_1^n r_i x_i \beta_i$. T^2 具有相同的特征向量, 但特征值为 $\{r_i^2\}$. 因此 $T^2\xi = \sum_1^n r_i^2 x_i \beta_i$. T^{-1} 存在当且仅当没有一个 $r_i = 0$, 这时它有相同的特征向量而特征值为 $\{1/r_i\}$. 因此 $T^{-1}\xi = \sum_1^n (x_i/r_i)\beta_i$. 如果 $P(t) = \sum_0^m a_n t^n$ 是任一个多项式, 则 $P(T)$ 将 β_i 带到 $P(r_i)\beta_i$. 因此 $P(T)\xi = \sum_0^m P(r_i)x_i\beta_i$.

到现在上面要表达的观点应该极其清楚了.

基中的法正交性更多的价值在上一节已经清楚了. 从根本上说, 它使我们能够用内积计算 ξ 的系数 $\{x_i\} : x_i = (\xi, \beta_i)$.

现在是关于在有限维理论中一般的特征值问题说上几句话的好时机. 是由于 T 的自伴性 (或矩阵 $\boldsymbol{t}$ 的对称性) 才使我们有可能做出上面完全的分析. 然而我们对 Hom$V$ 中任意一个 T 能说的远远不能令人满意.

我首先注意到 T 的特征值可以被代数地决定, 因为 λ 是一个特征值当且仅当 $T - \lambda I$ 不是单射, 或者等价地说, 是奇异的, 而我们知道 $T - \lambda I$ 为奇异当且仅当它的行列式 $\Delta(T - \lambda I)$ 为 0. 如果我们任选 V 的一个基, 则 $T - \lambda I$ 的行列式是它的矩阵 $\boldsymbol{t} - \lambda \boldsymbol{e}$ 的行列式, 我们以后在第七章的公式表明这是一个 λ 的 n 次多项式. 容易看出这个多项式与基无关, 它被称作 T 的*特征多项式*. 因此 T 的特征值正好是 T 的特征多项式的根.

但是 T 不必有任何特征向量! 例如考虑在笛卡儿平面中一个 90° 的旋转. 这是个映射 $T : \prec x, y \succ \mapsto \prec y, x \succ$. 因而 $T(\delta^1) = \delta^2$, $T(\delta^2) = -\delta^1$, 故 T 的矩阵是

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

于是 $T - \lambda$ 的矩阵为

$$\begin{bmatrix} \lambda & -1 \\ 1 & \lambda \end{bmatrix},$$

T 的特征多项式从而为此矩阵的行列式:$\lambda^2 + 1$. 由于此多项式在 $\mathbb{R}$ 上不可约, 所以没有特征值.

如果我们在复二维空间 $\mathbb{C}^2$ 上考虑具有相同矩阵的变换就会注意到结果有多么不同. 这里的纯量域为复数系, T 为映射 $\prec z_1, z_2 \succ \mapsto \prec -z_2, z_1 \succ : \mathbb{C}^2 \to \mathbb{C}^2$.

但现在 $\lambda^2+1=(\lambda-i)(\lambda+i)$, T 具有特征值 $\pm i$! 要求出 i 的特征向量, 我们需解方程 $T(z)=iz$, 即方程 $\prec -z_2,z_1\succ=\prec iz_1,iz_2\succ$, 或 $z_2=-iz_1$. 因此 $\prec 1,-i\succ$(或 $i\prec 1,-i\succ=\prec i,1\succ$) 是 i 的在纯量因子范围内惟一的特征向量.

回到我们的实的理论. 如果 $T\in \mathrm{Hom}V, n=d(V)$, 从而使 $d(\mathrm{Hom}V)=n^2$, 因此 n^2+1 个向量 $\{T^i\}_0^{n^2}$ 的集合在 $\mathrm{Hom}V$ 中线性相关. 这恰恰等于说对某个次数 $\leqslant n^2$ 的多项式 p 有 $p(T)=0$. 现假设 r 是 T 的一个特征值, 且 $T(\xi)=r\xi$. 于是 $p(T)(\xi)=p(r)\xi=0$, 所以 $p(r)=0$. 这就是说, T 的每个特征值是多项式 p 的一个根. 反之, 如果 $p(r)=0$, 我们则从代数的剩余定理知道, $t-r$ 是多项式 $p(t)$ 的一个因子, 因此 $(t-r)^m$ 将是 p 的互素因式中的一个. 现在设 p 是 T 的极小多项式 (见习题 3.5). 第一章的定理 5.5 告诉我们 $(T-rI)^m$ 在 V 的一个对应子空间 N 上为零, 因此特别地, $T-rI$ 在 N 上的限制不是单射. 就是说 r 是个特征值. 我们已经证明了:

定理 3.3 *T 的特征值是所有使 $p(T)=0$ 的多项式 $p(t)$ 的零点 (根), 并且正好是极小多项式的根.*

习题

3.1 用定义等式 $(T\xi,\eta)=(\xi,T\eta)$ 证明 $\mathrm{Hom}V$ 中所有自伴元的集合 SA 是个子空间. 类似地证明, 如果 S 和 T 是自伴的则 ST 为自伴当且仅当 $ST=TS$. 最后结论是, 如果 T 为自伴的则对任意多项式 p, $p(T)$ 也是自伴的.

3.2 证明如果 T 为自伴则 $S=T^2$ 为非负. 因此证明, 如果 T 为自伴, $a>0$, 则 T^2+aI 不可能是零变换.

3.3 设 $p(t)=t^2+bt+c$ 为一不可约二次多项式 (即 $b^2<4c$), 又设 T 为一个自伴变换. 证明 $p(T)\neq 0$.(配平方并应用前面的几个习题.)

3.4 设 T 为自伴和幂零的 (即对某个 n, $T^n=0$). 证明 $T=0$. 这可以各种方法加以证明. 方法之一是首先证明 $n=2$ 的情形, 然后由归纳法对 $n=2^m$ 证明. 最后对 n 包括在 2 的幂的情形:$2^m\leqslant n\leqslant 2^{m+1}$ 加以证明.

3.5 设 V 为任一向量空间, T 为 $\mathrm{Hom}V$ 中任一元. 假设存在多项式 q 使得 $q(T)=0$, 并设 p 为这样的具最小次数的多项式. 证明 p 是惟一的 (差一个常数因子的范围内). 称它为 T 的*极小多项式*.证明, 如果我们对 T 的极小多项式 p 应用第一章的定理 5.5, 则这些子空间 N_i 必定都是非平凡的.

3.6 代数基本定理的一个推论说, 实系数的一个多项式可以分解为线性因子 $(t-r)$ 和不可约的二次因子 (t^2+bt+c) 的乘积. 设 T 为一个有限维希尔伯特空间上的一个自伴变

换, $p(t)$ 为其极小多项式. 将前面的评注, 第一章定理 5.5 和习题 3.1~3.4 用来得到定理 3.1 的一个新的证明.

3.7 证明, 如果 T 是准希尔伯特空间 V 上一个自伴变换, 则其零空间是其值域空间的正交补:$N(T) = (R(T))^{\perp}$. 结论是, 如果 V 为希尔伯特空间, 则一个自伴 T 为单射当且仅当它的值域稠于 V.

3.8 假定上面的习题, 证明如果 V 为希尔伯特空间, T 为 Hom$V$ 中的自伴元并且它为下有界的 (同时也有界), 则 T 为满射.

3.9 设 T 为自伴且非负, 并令 $m = \text{lub}\{(T\xi, \xi) : \|\xi\| = 1\}$. 用施瓦兹不等式和引理 3.2 的不等式证明 $m = \|T\|$.

3.10 设 V 为希尔伯特空间, T 为 Hom$V$ 中的自伴元, 并设 $m = \text{lub}\{(T\xi, \xi) : \|\xi\| = 1\}$. 证明如果 $a > m$, 则 $a - T(= aI - T)$ 可逆; 且 $\|(a-T)^{-1}\| \leqslant 1/(a-m)$. (应用施瓦兹不等式, m 的定义和习题 3.8.)

3.11 设 P 是准希尔伯特空间 V 上的一个有界线性变换, 它是在第一章的意义下的一个投影. 证明如果 P 是自伴的, 则 P 是个正交投影.现证明其逆结论.

3.12 设 V 为有限维希尔伯特空间, T 为 Hom$V$ 中的自伴, 并假设在 Hom$V$ 中 S 与 T 可交换. 证明定理 3.1 和引理 3.3 中的子空间 M_i 在 S 下不变.

3.13 称有限维希尔伯特空间 V 上的一个自伴变换 T 有一个单谱是说, 如果它的所有特征值互不相同. 由此我们知道所有子空间 M_i 都是一维的. 假设 T 是一个具有单谱的自伴变换, 并设 S 与 T 可交换. 证明 S 也为自伴.(应用上一个习题.)

3.14 设 H 为希尔伯特空间, 并设 $\omega[\xi, \eta]$ 为在 $H \times H$ 上的有界双线性形式. 就是说, 存在一个常数 b 使得

$$|\omega[\xi, \eta]| \leqslant b\|\xi\| \ \|\eta\|$$

对所有 $\xi, \eta \in H$ 成立. 证明在 Hom$V$ 中有一个惟一的 T 使得 $\omega[\xi, \eta] = (\xi, T\eta)$. 证明 T 为自伴当且仅当 ω 为对称.

5.4 正交变换

假定 V 是希尔伯特空间, 并因此 $\theta : V \to V^*$ 是一个同构, 我们自然把任何 $T \in \text{Hom}V$ 的伴随 $T^* \in \text{Hom}V^*$ 换为对应的变换 $\theta^{-1} \circ T^* \circ \theta \in \text{Hom}V$. 在希尔伯特空间理论中正是这个映射被称作 T 的伴随, 并以 T^* 表示. 于是, 完全像在自伴 T 的我们的讨论那样, 我们看出, 对所有 $\alpha, \beta \in V$

$$(T\alpha, \beta) = (\alpha, T^*\beta),$$

并且 T^* 由此等式惟一决定. 最后, T 为自伴当且仅当 $T=T^*$.

虽然它确实等于把 T^* 引进 Hom$V$ 的上述方法, 我们还是要做一个直接的定义如下. 对每个 η, 映射 $\xi\mapsto(T\xi,\eta)$ 为线性有界, 故为 V^* 中一个元, 并由定理 2.4 知, 它由 V 中一个元 β_η 按公式 $(T\xi,\eta)=(\xi,\beta_\eta)$ 给出. 现我们验证 $\eta\mapsto\beta_\eta$ 是线性和有界的, 从而是 Hom$V$ 中的一个元, 我们称其为 T^*, 等等.

引理 3.1 的矩阵计算逐字逐句地推广以证明 T^* 在 Hom$V$ 中的矩阵为 T 的矩阵 t 的转置 t^* .

在希尔伯特空间上另一个非常重要类型的变换是保持由积的变换.

定义 4.1 称一个变换 $T\in\mathrm{Hom}V$ 是正交的是说如果

$$(T\alpha,T\beta)=(\alpha,\beta)\text{对所有}\alpha,\beta\in V\text{成立}.$$

由上面的基本伴随等式知道, 这完全等价于 $(\alpha,T^*T\beta)=(\alpha,\beta)$ 对所有 $\alpha,\beta\in V$ 成立. 因而 $T^*T=I$. 因为对正交 T 有 $\|T\alpha\|^2=\|\alpha\|^2$, 故一个正交变换 T 为单射, 并且如果 V 是有限维的, 它因而也是可逆的. 不管 V 是否有限, 如果 T是可逆, 则上面的条件就成了 $T^*=T^{-1}$.

如果 $T\in\mathrm{Hom}\mathbb{R}^n$, 方程 $T^*T=I$ 的矩阵形式自然为 $t^*t=e$, 如果把这个写出来它成为, 对所有 i,j 有

$$\sum_{k=1}^n t_{ki}t_{kj}=\delta_j^i,$$

它不过是说 t 的列构成了 $\mathbb{R}^n$ 中一个法正交集 (从而是个基). 我们有

定理 4.1 一个变换 $T\in\mathrm{Hom}\mathbb{R}^n$ 是正交的当且仅当标准基 $\{\delta^i\}_1^n$ 在 T 下的像是另一个基 (关于标准内积的).

当然这个条件的必要性显然来自保持正交性定义的内积, 而充分性也可用 α 和 β 的基展开式直接验证.

我们现在以不同的术语来叙述特征基定理. 我们所说的对角矩阵是指一个除主对角线外处处为零的矩阵.

定理 4.2 设 $t=\{t_{ij}\}$ 为一对称 $n\times n$ 矩阵. 则存在一个正交 $n\times n$ 矩阵 b 使得 $b^{-1}tb$ 为一对角矩阵.

证明 因为由 t 定义的变换 $T\in\mathrm{Hom}\mathbb{R}^n$ 是自伴的, 故存在一个法正交基 $\{b^i\}_1^n$, 由 T 的特征向量组成, 其相应的特征值 $\{r_i\}_1^n$. 设 B 为由 $B(\delta^j)=b^j$ 定

义的正交变换, $j=1,\cdots,n$.(这里 n 联组 $\boldsymbol{b}^j$ 是 B 的矩阵 $\boldsymbol{b}=(b_{ij})$ 的列.) 于是 $(B^{-1}\circ T\circ B)(\delta^j)=r_j\delta^j$, 由于 $(B^{-1}\circ T\circ B)(\delta^j)$ 是矩阵 $\boldsymbol{b}^{-1}\boldsymbol{tb}$ 的第 j 列, 我们看出 $\boldsymbol{s}=\boldsymbol{b}^{-1}\boldsymbol{tb}$ 为对角的, 其中 $s_{jj}=r_j$. □

为了后面的应用我们还要下面的结果.

定理 4.3 *在有限维希尔伯特空间 V 上的一个可逆元 $T\in\mathrm{Hom}V$ 可以表示为形式 $T=RS$, 其中 R 为正交而 S 为自伴和正定的.*

证明 对任意 T, 因为 $(T^*T)^*=T^*T^{**}=T^*T$ 故 T^*T 为自伴. 令 $\{\varphi_i\}_1^n$ 为一个法正交的特征基, $\{r_i\}_1^n$ 为相应的 T^*T 的特征值. 于是 $0<\|T\varphi_i\|^2=(T^*T\varphi_i,\varphi_i)=(r_i\varphi_i,\varphi_i)=r_i$, 对每个 i 成立. 因为 T^*T 的所有特征值为正, 我们可以定义一个正平方根 $S=(T^*T)^{1/2}$ 为 $S\varphi_i=(r_i)^{1/2}\varphi_i, i=1,2,\cdots,n$. 可以清楚看到 $S^2=T^*T$, 并且 S 为自伴的.

因为有 $(ST^{-1}\alpha,ST^{-1}\beta)=(T^{-1}\alpha,S^2T^{-1}\beta)=(T^{-1}\alpha,T^*TT^{-1}\beta)=(T^{-1}\alpha,T^*\beta)=(TT^{-1}\alpha,\beta)=(\alpha,\beta)$, 故 $A=ST^{-1}$ 是正交的. 因为 $T=A^{-1}S$, 令 $R=A^{-1}$, 则得到定理. □

不难看出上面 T 的分解是惟一的. 又从 TT^* 出发, 我们可将 T 表示为形式 $T=SR$, 其中 S 为自伴和正定而 R 为正交.

我们称这些因子分解为 T 的极分解, 这是因为它们的作用多少有些像复数的极坐标分解 $z=re^{i\theta}$.

推论 4.1 *任一非异 $n\times n$ 矩阵 $\boldsymbol{t}$ 可以表示为 $\boldsymbol{t}=\boldsymbol{udv}$, 其中 $\boldsymbol{u}$ 和 $\boldsymbol{v}$ 为正交, $\boldsymbol{d}$ 为对角.*

证明 由此定理我们有 $\boldsymbol{t}=\boldsymbol{rs}$, 其中 $\boldsymbol{r}$ 正交, $\boldsymbol{s}$ 为对称. 由定理 4.2, $\boldsymbol{s}=\boldsymbol{bdb}^{-1}$, $\boldsymbol{d}$ 为对角, $\boldsymbol{b}$ 为正交, 因此 $\boldsymbol{t}=\boldsymbol{rs}=(\boldsymbol{rb})\boldsymbol{db}^{-1}=\boldsymbol{udv}$, 其中 $\boldsymbol{u}=\boldsymbol{rb}$, $\boldsymbol{u}=\boldsymbol{b}^{-1}$ 都为正交. □

习题

4.1 设 V 为希尔伯特空间, 并设 S 和 $T\in\mathrm{Hom}V$ 满足

$$(T\xi,\eta)=(\xi,S\eta)\quad \text{对所有 } \xi,\eta \text{ 成立.}$$

写出恒等式 $S=\theta^{-1}\circ T^*\circ\theta$ 的证明.

4.2 写出类比于引理 3.1 的证明以表明 T^* 的矩阵是 T 的矩阵的转置.

4.3 再一次证明如果 $(\xi,\eta)=(\xi,\zeta)$ 对所有 ξ 成立, 则 $\eta=\zeta$. 结论是, 如果 $S,T\in \mathrm{Hom}V$ 使得 $(\xi,T\eta)=(\xi,S\eta)$ 对所有 η 成立, 则 $T=S$.

4.4 如果 $\{\boldsymbol{a},\boldsymbol{b}\}$ 为 $\mathbb{R}^2$ 中的法正交基, 并设 $\boldsymbol{t}$ 为其列为 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 的 2×2 矩阵. 以直接计算证明 $\boldsymbol{t}$ 的行也是法正交的.

4.5 再次陈述如果 V 为有限维, 且如果 $S,T\in \mathrm{Hom}V$ 满足 $S\circ T=I$, 则 T 可逆, 并有 $S=T^{-1}$ 这个论断的理由. 现设 V 为有限维希尔伯特空间, 并设 T 为 $\mathrm{Hom}V$ 中的正交变换. 证明 T^* 也是正交的.

4.6 设 $\boldsymbol{t}$ 为一 $n\times n$ 的矩阵, 其列构成 $\mathbb{R}^n$ 的一个正交基. 证明 $\boldsymbol{t}$ 的行也构成一个正交基.(用上一个习题.)

4.7 证明非负自伴变换 S 在一个有限维希尔伯特空间上有一个惟一决定的自伴平方根.

4.8 证明如果 V 是有限维希尔伯特空间, $T\in \mathrm{Hom}V$, 则定理 4.3 中 T 的 "极分解"$T=RS$ 是惟一的.(应用上面习题.)

5.5 紧变换

当 V 是无穷维希尔伯特空间时定理 3.1 不再有效了. 一个自伴变换 T 一般地不具有足够多的特征向量以形成 V 的一个基, 需要一个更为精细的分析以便处理 "连续谱" 还有 "离散谱" 的情况. 这种更为深刻和丰富的情形是需要在研究生水平上进一步学习希尔伯特空间理论的理由, 这也是量子力学的数学结构复杂性的一个源头.

但是, 有一个非常重要的特殊情形, 在那里特征基定理是有效的, 并在下一章中有着令人吃惊的应用.

定义 5.1 *设 V 和 W 为任意赋范线性空间, S 为 V 中的单位球. $\mathrm{Hom}(V,W)$ 中的一个变换 T 被称作紧的是说 $T[S]$ 在 W 中的闭包是列紧的.*

定理 5.1 *设 V 是任一准希尔伯特空间, $T\in \mathrm{Hom}V$ 为自伴和紧的. 则准希尔伯特空间 $R=T$ 的值域有一个法正交基 $\{\varphi_i\}$, 它完全由 T 的特征向量组成, 而且相应的特征值的序列 $\{r_n\}$ 收敛于 0(或者有限).*

证明 除了有一点不同的开头以外, 这个定理的证明正好与定理 3.1 的证明相同. 令 $m=\|T\|=\mathrm{lub}\{\|T(\xi)\|:\|\xi\|=1\}$, 并选取序列 $\{\xi_n\}$ 使 $\|\xi_n\|=1$ 对所

有 n 成立, 同时 $\|T(\xi_n)\| \to m$. 于是

$$((m^2 - T^2)\xi_n, \xi_n) = m^2 - \|T(\xi_n)\|^2 \to 0,$$

并且因为 $m^2 - T^2$ 是个非负的自伴变换, 引理 3.2 便告诉了我们 $(m^2 - T^2)(\xi_n) \to 0$. 但是因为 T 为紧, 我们可以假定 (如有必要可转而用子序列)$\{T(\xi_n)\}$ 收敛于譬如 β. 于是 $T^2\xi_n \to T\beta$, 故 $m^2\xi_n \to T\beta$. 因此 $\xi_n \to T\beta/m^2$ 而 $\beta = \lim T\xi_n = T^2(\beta)/m^2$. 因为 $\|\beta\| = \lim\|T(\xi_n)\| = m$, 所以我有了一个非零向量 β 使得 $T^2(\beta) = m^2\beta$. 令 $\alpha = \beta/\|\beta\|$.

于是我们已经找到了一个向量 α, 满足 $\|\alpha\| = 1$ 和 $0 = (m^2 - T^2)(\alpha) = (m - T)(m + T)(\alpha)$. 那么或是 $(m + T)(\alpha) = 0$, 这时 $T(\alpha) = -m\alpha$, 或是 $(m + T)(\alpha) = \gamma \neq 0$ 而 $(m - T)\gamma = 0$, 这时 $T\gamma = m\gamma$. 因此总存在一个向量 $\varphi_1(\alpha$ 或 $\gamma/\|\gamma\|)$ 使得 $\|\varphi_1\| = 1$ 而 $T(\varphi_1) = r_1\varphi_1$, 其中 $|r_1| = m$. 我们现在像在定理 3.1 中同样地进行下去.

为了保持记号的一致性, 我们令 $m_1 = m, V_1 = V$, 现令 $V_2 = \{\varphi_1\}^{\perp}$. 于是 $T[V_2] \subset V_2$, 这是因为如果 $\alpha \perp \varphi_1$, 则 $(T\alpha, \varphi_1) = (\alpha, T\varphi_2) = r_1(\alpha, \varphi_1) = 0$. 因此 $T \restriction V_2$ 为紧且自伴, 并且如果 $m_2 = \|T \restriction V_2\|$, 则存在 φ_2, $\|\varphi_2\| = 1$, 及 $T(\varphi_2) = r_2\varphi_2$, 其中 $|r_2| = m_2$. 归纳地继续下去, 得到了一个法正交序列 $\{\varphi_n\} \subset V$ 及序列 $\{r_n\} \subset \mathbb{R}$ 使得 $T\varphi_n = r_n\varphi_n, |r_n| = \|T \restriction V_n\|$, 其中

$$V_n = \{\varphi_1, \cdots, \varphi_{n-1}\}^{\perp}.$$

由于这是最感兴趣的情形, 我们姑且假定对所有的 $n, r_n \neq 0$. 于是我们断言 $|r_n| \to 0$. 因为无论怎样, $|r_n|$ 是递降的, 如果它不趋向 0, 则存在一个 $b > 0$ 使 $|r_n| \geqslant b$ 对所有 n 成立. 那么 $\|T(\varphi_i) - T(\varphi_j)\|^2 = \|r_i\varphi_i - r_j\varphi_j\|^2 = \|r_i\varphi_i\|^2 + \|r_j\varphi_j\|^2 = r_i^2 + r_j^2 \geqslant 2b^2$, 对所有 $i \neq j$ 成立, 从而 $\{T(\varphi_i)\}$ 便不能有收敛的子序列, 这与 T 的紧性相矛盾. 因此 $|r_n| \downarrow 0$.

最后, 我们必须证明法正交序列 $\{\varphi_i\}$ 是 R 的一个基. 如果 $\beta = T(\alpha)$, 并且如果 $\{b_n\}$ 和 $\{a_n\}$ 分别为 β 和 α 的傅里叶系数, 我们于是期望有 $b_n = r_n a_n$, 但是这容易得到验证: $b_n = (\beta, \varphi_n) = (T(\alpha), \varphi_m) = (\alpha, T(\varphi_n)) = (\alpha, r_n, \varphi_n) = r_n(\alpha, \varphi_n) = r_n a_n$. 这恰好是说 $T(a_n\varphi_n) = b_n\varphi_n$, 因而 $\beta - \sum_1^n b_i\varphi_i = T(\alpha - \sum_1^n a_i\varphi_i)$. 现已知 $\alpha - \sum_1^n a_i\varphi_i$ 正交于 $\{\varphi_i\}_1^n$, 因此是 V_{n+1} 中一个元素, 并且 T 在 V_{n+1} 上的范数是 $|r_{n+1}|$. 又由毕达哥拉斯定理, $\|\alpha - \sum_1^n a_i\varphi_i\| \leqslant \|\alpha\|$. 总合起来, 我们最后有

$$\|\beta - \sum_1^n b_i\varphi_i\| \leqslant |r_{n+1}| \cdot \|\alpha\|,$$

并且因为 $r_{n+1} \to 0$, 这表明 $\beta = \sum_1^\infty b_i\varphi_i$. 故 $\{\varphi_i\}$ 是 $R(T)$ 的一个基. 由于 T 是自伴的, 也有 $N(T) = R(T)^\perp = \{\varphi_i\}^\perp = \bigcap_1^\infty V_i$.

如果有某个 $r_i = 0$, 则有一个使 $r_n = 0$ 的第一个 n. 这时 $\|T \restriction V_n\| = |r_n| = 0$, 故 $V_n \subset N(T)$. 但因为 $\varphi_i = T(\varphi_i)/r_i$, 如果 $i < n$ 则有 $\varphi_i \in R(T)$, 因而 $N(T) = R(T)^\perp \subset \{\varphi_1, \cdots, \varphi_{n-1}\}^\perp = V_n$. 因此, $N(T) = V_n$, $R(T)$ 由 $\{\varphi_i\}_1^{n-1}$ 张成. □

第六章　微分方程

这一章并非是一本小的微分方程的教科书; 我们留下未讲太多了. 尽管我们也要讲一两个实际问题, 但我们主要关心的是这门学科的理论中几个方面. 我们的第一个目标是常微分方程的基本存在性和惟一性定理, 我们把它们作为不动点定理的漂亮应用而加以证明. 其次, 我们将考察线性理论, 在那里我们对最前面两章素材做出关键性的应用, 并对实际求解过程得到相当详细的阐述. 就我们阐述的范围而言所关联的是*初值*问题, 它与在包含解曲线空间中, 通过某些最初描述的点的惟一解的存在性, 以及在某些场合下的求解方式有关. 然而, 此学科的一些最重要的方面关联着那些被称作*边值*问题的东西, 而我们最终和最深层次的努力所对准的目标是朝着这个博大的领域迈出第一步. 这使我们直接牵涉到第五章的理论, 因为我们会发现自己正在研究自伴算子. 事实上, 关于傅里叶级数展开式的基本定理就来自于对一个微分算子的右逆元为紧自伴算子的认识.

6.1　基本定理

设 A 为巴拿赫空间 W 的一个开子集, 设 I 为 $\mathbb{R}$ 中的一个开区间, 并设 $F: I \times A \to W$ 连续. 我们要研究微分方程

$$d\alpha/dt = F(t, \alpha).$$

这个方程的一个解是个函数 $f: J \to A$, 其中 J 为 I 的一个开子区间, 使得 $f'(t)$ 存在且对每个 J 中的 t 有

$$f'(t) = F(t, f(t)).$$

注意, 一个解 f 必须是连续可微的, 因为 f' 的存在性意味着 f 的连续性, 于是 $f'(t) = F(t, f(t))$ 由 F 的连续性是连续的.

我们将看到如果 F 具有连续的第二个偏微分, 则存在一个惟一确定的通过任意点

$$\prec t_0, \alpha_0 \succ \in I \times A$$

的"局部"解. 说解 f 通过 $\prec t_0, \alpha_0 \succ$ 我们的意思自然是说 $\alpha_0 = f(t_0)$. 要求解 f 当 $t = t_0$ 时有值 α_0 被称作一个*初值条件*.

$dF^2_{\prec t,\alpha \succ}$ 的存在性和连续性通过中值定理表明 $F(t, \alpha)$ 对 α 是*局部地一致利普希茨的*. 我们的意思是对任意点 $\prec t_0, \alpha_0 \succ \in I \times A$, 存在一个邻域 $M \times N$ 和一个常数 b 使得 $\|F(t, \xi) - F(t, \eta)\| \leqslant b\|\xi - \eta\|$ 对 M 中所有的 t 和 N 中所有的 ξ, η 成立. 要明白这点, 我们简单地选取在 t_0 和 α_0 的球 M 和 N 使得 $dF^2_{\prec t,\alpha \succ}$ 在 $M \times N$ 上以 b 为界, 并应用第三章的定理 7.4. 这就是我们在下面实际使用的条件.

定理 1.1 *设 A 为一个巴拿赫空间 W 的一个开子集, I 为 $\mathbb{R}$ 中的一个开区间, 并设 F 是从 $I \times A$ 到 W 的一个连续映射, 它对它的第二个变量为局部地一致利普希茨连续. 于是对 $I \times A$ 中任意点 $\prec t_0, \alpha_0 \succ$, 对 α_0 的某个邻域 U 和对任意包含 t_0 的充分小的区间 J, 存在一个从 J 到 U 的惟一函数 f, 它是此微分方程的通过点 $\prec t_0, \alpha_0 \succ$ 的解.*

证明 如果 f 是 J 上通过 $\prec t_0, \alpha_0 \succ$ 的一个解, 于是积分给出

$$f(t) - f(t_0) = \int_{t_0}^{t} F(s, f(s))ds,$$

故对 $t \in J$,

$$f(t) = \alpha_0 + \int_{t_0}^{t} F(s, f(s))ds.$$

反之, 如果 f 满足这个"积分方程", 则微积分的基本定理蕴涵了在 J 上 $f'(t)$ 存在并等于 $F(t, f(t))$, 从而 f 是此微分方程的解并虽然它通过 $\prec t_0, \alpha_0 \succ$. 现在对

于任意连续函数 $f: J \to A$ 我们可定义 $g: J \to W$ 为

$$g(t) = \alpha_0 + \int_{t_0}^{t} F(s, f(s))ds,$$

而我们上面的论证表明 f 是此微分方程的解当且仅当 f 是映射 $K: f \mapsto g$ 的不动点. 这提示我们要试图证明 K 是个压缩映射, 从而我们可以应用不动点定理.

我们从选取 $\prec t_0, \alpha_0 \succ$ 的一个邻域 $L \times U$ 开始, 它使 $F(t, \alpha)$ 在它上面有界且对 α 在 t 上为一致地利普希茨连续. 设 J 为 L 中包含 t_0 的某个开子区间, 并设 $\mathfrak{BC}(J, W)$ 为从 J 到 W 的有界连续函数的巴拿赫空间. 我们在后面的计算将表明我们必须取 J 如何小. 我们假定邻域 U 是在 α_0 的一个球, 半径为 r, 同时我们也考虑 V 中的函数 $\mathrm{u} = B_r(\overline{\alpha}_0)$ 的球, 其中 $\overline{\alpha}_0$ 是取值 α_0 的常值函数. 于是在 u 中的任意 f 的值域在 U 中, 使 $F(t, f(t))$ 是确定的, 有界的和连续的. 那就是说, 前面所定义的 K 将球 u 映射到 V 内.

现在进行计算. 设 F 在 $L \times U$ 上以 m 有界, 并设 δ 为 J 的长度. 于是由对积分的范数不等式 (见第四章第 10 节),

$$\|K(\overline{\alpha}_0) - \overline{\alpha}_0\|_\infty = \mathrm{lub}\left\{\left\|\int_{t_0}^{t} F(s, \alpha_0)ds\right\| : t \in J\right\} \leqslant \delta m \tag{1}$$

又, 如果 f_1 和 f_2 在 u 中, 并且如果 e 是 F 在 $L \times U$ 上的一个利普希茨常数, 则

$$\begin{aligned}\|K(f_1) - K(f_2)\|_\infty &= \mathrm{lub}\left\{\left\|\int_{t_0}^{t} F(s, f_1(s)) - F(s, f_2(s))\right\|\right\} \\ &\leqslant \delta \mathrm{lub}\{\|F(s, f_1(s)) - F(s, f_2(s))\|\} \\ &\leqslant \delta c \mathrm{lub}\{\|f_1(s) - f_2(s)\|\} \\ &\leqslant \delta c \|f_1 - f_2\|_\infty. \end{aligned} \tag{2}$$

由 (2) 式我们看出 K 在 $\delta c < 1$ 时是具有常数 $C = \delta c$ 的一个压缩映射, 并且由 (1) 式我们看出 K 移动球 u 的中心 $\overline{\alpha}_0$ 的距离在 $\delta m < (1 - \delta c)r$ 时小于 $(1 - C)r$. 这个对 δ 的双重要求等价于

$$\delta < \frac{r}{m + cr},$$

从而以任意这样的 δ, 此定理由不动点定理的一个推论得到 (第四章定理 9.1 的推论 2.) □

推论 1.1 如果 $F: I \times A \to W$ 连续且有一个连续的第二个偏微分, 则上述定理成立.

我们下面要证明任意两个通过 $\prec t_0, \alpha_0 \succ$ 必定在它们定义域的相交部分上相同 (在定理 1.1 的假设条件下).

引理 1.1 设 g_1 和 g_2 是 $d\alpha/dt = F(t,\alpha)$ 通过 $\prec t_0, \alpha_0 \succ$ 的任意两个解. 则对所有在它们定义域的交 $J = J_1 \cap J_2$ 上的所有 t 有 $g_1(t) = g_2(t)$.

证明 如果上述结论不真则存在一点 $s \in J$ 使 $g_1(s) \neq g_2(s)$. 假设 $s > t_0$, 并令 $C = \{t : t > t_0$ 且 $g_1(t) \neq g_2(t)\}$ 以及 $x = \text{glb}\ C$. 集合 C 由于 g_1 和 g_2 是连续而为开, 从而 x 不在 C 中, 也就是说, $g_1(x) = g_2(x)$. 称此公共值为 α, 并将上面的定理用于 $\prec x, \alpha \succ$. 取 r 使得 $B_r(\alpha) \subset A$, 对此我们选取 δ 足够小, 使得此微分方程有一个经过 $\prec x, \alpha \succ$ 的从 $(x-\delta, x+\delta)$ 到 $B_r(\alpha)$ 的惟一的解 g, 我们也取 δ 足够小使得 g_1 和 g_2 在 $(x-\delta, x+\delta)$ 上的限制的值域在 $B_r(\alpha)$ 中. 那么由 g 在此区间上的惟一性有 $g_1 = g_2 = g$, 但这与 x 的定义相矛盾. 因此在它们的定义域的交有 $g_1 = g_2$. □

这个引理证明我们能够把定理中加在 f 值域的限制去掉.

定理 1.2 设 A, I, F 同于定理 1.1 中的假定. 于是对 $I \times A$ 中任意点 $\prec t_0, \alpha_0 \succ$ 和 t_0 的充分小的任意区间邻域 J, 存在一个从 J 到 A 的通过 $\prec t_0, \alpha_0 \succ$ 的惟一解.

整体解 我们对微分方程 $d\alpha/dt = F(t,\alpha)$ 已经找到的这些解只在初值点 t_0 的充分小邻域中有定义, 并因此而被称作局部解. 现在我们离开到这种局部解的端点附近的一个点 $\prec t_1, \alpha_1 \succ$, 然后考虑在 $\prec t_1, \alpha_1 \succ$ 周围的局部解, 首先它必定与我们第一个解在两个定义域的相交部分一致, 其次它一般地将超越 t_1 扩张到比第一个解更远, 从而这两个局部解将合并起来做成一个在比原来各自分别给出的区间更大的 t 区间上的一个解. 我们可以按此方式不断扩张我们原来的解到那个被称作整体解的解, 这是由拼接局部解得到了. 这些像上面描述的概念多少有些含糊不清, 我们现在转到整体解的一个更为精确的构造.

给出 $\prec t_0, \alpha_0 \succ \in I \times A$, 说 $\mathfrak{F}$ 为所有通过 $\prec t_0, \alpha_0 \succ$ 的所有解的族. 因此 $g \in \mathfrak{F}$ 当且仅当 g 是在一个区间 $J \subset I$ 上的解, 有 $t_0 \in J$ 且 $g(t_0) = \alpha_0$. 引理 1.1 完全表明 $\mathfrak{F}$ 中所有函数 g 的并[1] f 本身也是这样的函数, 因为如果 $\prec t_1, \alpha_1 \succ \in g_1$ 和 $\prec t_2, \alpha_2 \succ \in g_2$, 则 $\alpha_1 = g_1(t) = g_2(t) = \alpha_2$.

又, f 是一个解, 这是因为在其定义域中任一点 x 的附近 f 与某个 $g \in \mathfrak{F}$ 一

[1] 记住我们是在取一个作为有序偶对的集合的函数, 使得这个函数族的并有确切的意义.

致. 按 f 的定义方式我们看出 f 是经过 $\prec t_0, \alpha_0 \succ$ 的惟一的极大解. 我们已经证明了下面的定理.

定理 1.3 设 $F: I \times A \to W$ 是个满足定理 1.1 的假设条件的函数. 于是通过每个 $\prec t_0, \alpha_0 \succ \in I \times A$ 存在微分方程 $d\alpha/dt = F(t, \alpha)$ 的一个惟一确定的极大解.

一般地, 我们总期望一个 "跑到 A 的边界里的" 极大解, 因而有一个真正含于 I 中的定义区间 J, 这就像图 6.1 所提示的那样.

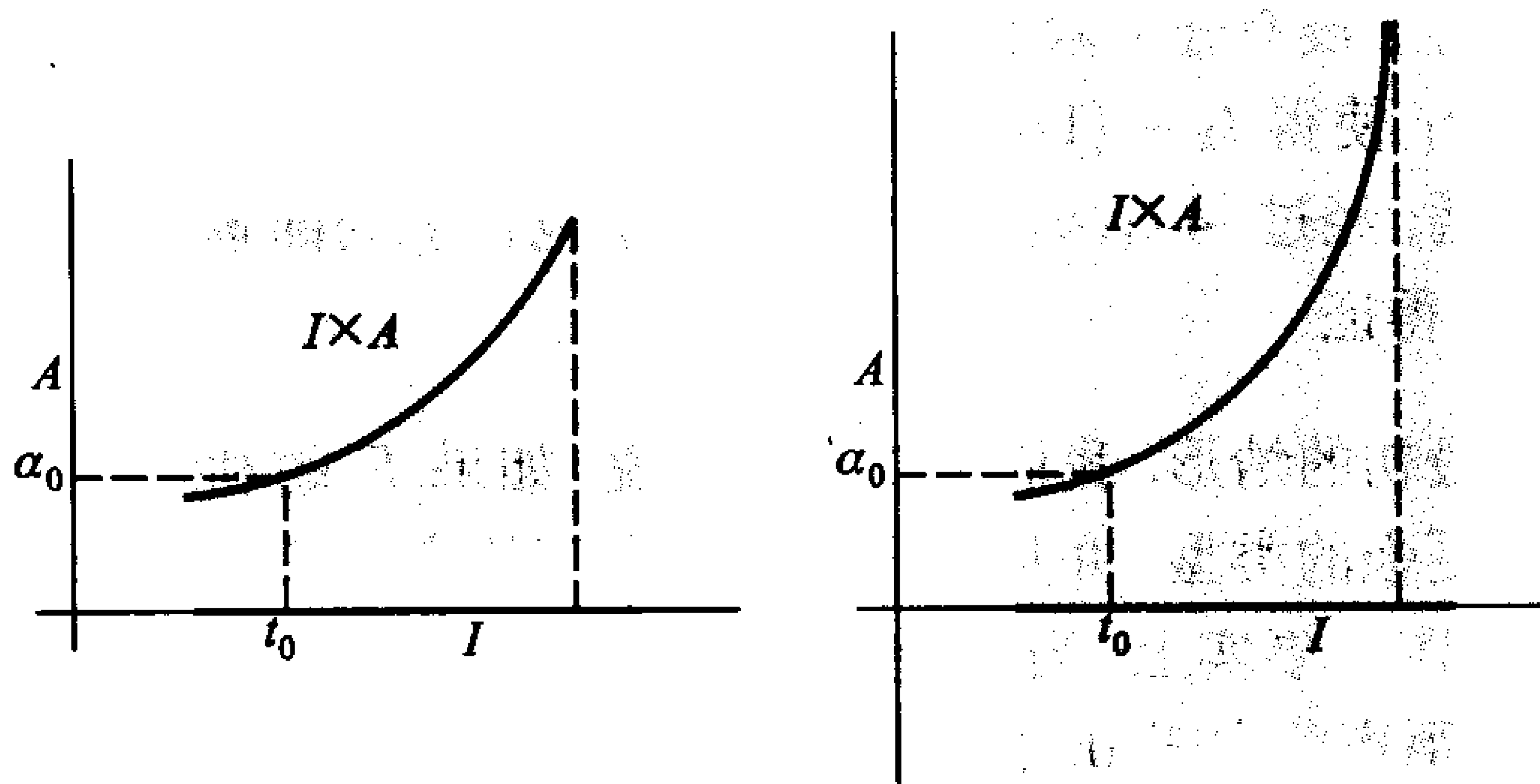


图 6.1

但是, 如果 A 是整个空间 W, 并且如果 $F(t, \alpha)$ 对每个 t 而言是 α 的利普希茨连续函数, 它具有的利普希茨界 $c(t)$ 为 t 的连续函数, 我们则可以证明每个极大解都是在整个 I 上的. 我们稍后就能看到这个条件对于线性方程而言是自然的.

定理 1.4 设 W 为巴拿赫空间, I 为 $\mathbb{R}$ 中一个开区间. 设 $F: I \times W \to W$ 为连续, 并假定存在连续函数 $c: I \to \mathbb{R}$ 使得对所有 $t \in I$ 和所有 $\alpha_1, \alpha_2 \in W$ 有

$$\|F(t, \alpha_1) - F(t, \alpha_2)\| \leqslant c(t)\|\alpha_1 - \alpha_2\|.$$

于是微分方程 $d\alpha/dt = F(t, \alpha)$ 的每个极大解都以整个 I 为其定义域.

证明 设若相反, 假设 g 是个极大解, 它的定义区间 J 的右端点 b 小于 I 的右端点. 我们选取一个有限开区间 L, 它包含了 b 并使 $\overline{L} \subset I$(见图 6.2). 因 $\overline{L}$ 为紧, 故连续函数 $c(t)$ 在 $\overline{L}$ 上有极大值 c. 在 $L \cap J$ 中选取任一点 t_1 十分靠近

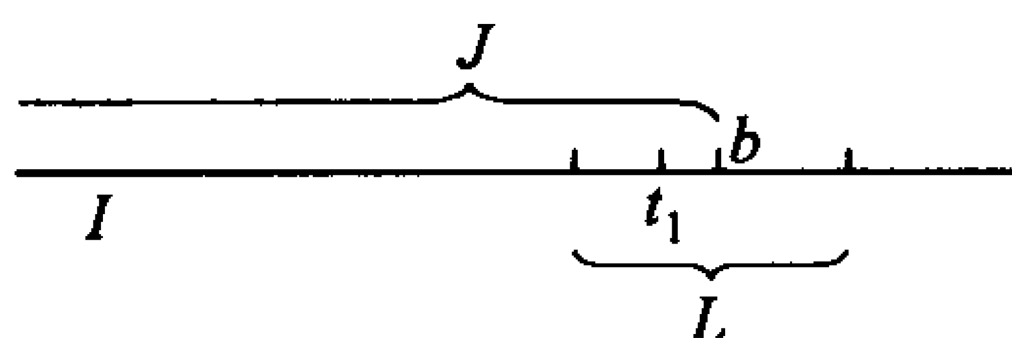


图 6.2

b 使得 $b-t_1<1/c$, 我们令 $\alpha_1=g(t_1)$ 及在 $\overline{L}$ 上 $m=\max\|F(t,\alpha_1)\|$. 以这些值 c,m 以及任意的 r, 定理 1.1 的证明给了我们一个通过 $\prec t_1,\alpha_1\succ$, 定义域为 $(t_1-\delta,t_1+\delta)$ 的局部解 f, 其中 δ 为小于 $r/(m+rc)=1/(c+(m/r))$ 的任意数. 由于现在我们对 r 没有任何限制 (因为 $A=W$), 这个对 δ 的界变成了 $1/c$, 而且由于我们选取 t_1 使得 $t_1+(1/c)>b$, 我们现在可以选取 δ 使得 $t_1+\delta>b$. 但是这导出了矛盾; 经过 $\prec t_1,\alpha_1\succ$ 的极大解 g 包含了局部解 f, 故而特别地, 有 $t_1+\delta\leqslant b$. 定理得证 . □

回到我们最初的情形, 我们可以下结论说, 如果 F 的利普希茨控制具有上面给出的那种更强的类型, 并且如果某个极大解的定义域 J 小于 I, 则开集 A 便不能是整个 W. 事实上, 当 t 趋向于作为 I 的内点的 J 的端点 b 时, $g(t)$ 到 A 的边缘的距离趋向于 0, 这是成立的. 那就是说, 现在它是个定理: 当 $t\to b,\rho(f(t),A')\to 0$. 它的证明比我们前面的论证更加复杂, 我们把它作为一组习题留给有兴趣的读者 .

n 阶方程　设 $A_1,A_2,\cdots,A_n$ 为一巴拿赫空间 W 的开子集, 设 I 为 $\mathbb{R}$ 中的开区间, $G:I\times A_1\times A_2\times\cdots\times A_n\to W$ 为连续. 我们考虑微分方程

$$d^n\alpha/dt^n=G(t,\alpha,d\alpha/dt,\cdots,d^{n-1}\alpha/dt^{n-1})\ .$$

一个函数 $f:J\to W$ 是这个方程的一个解指的是, 如果 J 是 I 的一个开子区间, f 在 J 上具有直到 n 阶的连续导数, 且 $f^{(i-1)}[J]\subset A_i, i=1,\cdots,n$, 使得

$$f^{(n)}(t)=G(t,f(t),f'(t),\cdots,f^{(n-1)}(t))$$

对所有 $t\in J$ 成立. 初值条件现在由一个点给出:

$$\prec t_0,\beta_1,\beta_2,\cdots,\beta_n\succ\in I\times A_1\times\cdots\times A_n\ .$$

它的基本定理几乎与前面的一样. 为简化记号, 设 $\boldsymbol{\alpha}$ 为 n 联组 $\prec\alpha_1,\alpha_2,\cdots,\alpha_n\succ\in W^n=V$, 令 $\boldsymbol{A}=\prod_1^n A_i$. 也设 ψ 为映射 $f\mapsto\prec f,f^1,\cdots,f^{(n-1)}\succ$. 于是解的方程成了 $f^{(n)}(t)=G(t,\psi f(t))$.

定理 1.5 设 $G: I \times \boldsymbol{A} \to W$ 如上, 并假设另外有 $G(t, \boldsymbol{\alpha})$ 对 $\boldsymbol{\alpha}$ 是局部一致的利普希茨连续. 于是对任意 $\prec t_0, \boldsymbol{\beta} \succ \in I \times \boldsymbol{A}$ 和对任意包含了 t_0 的充分小的开区间 J, 存在一个从 J 到 W 的惟一函数 f 使得 f 是上面的 n 阶方程的解, 并满足初值条件 $\psi f(t_0) = \boldsymbol{\beta}$.

证明 有一个古老而标准的手段来把单个的 n 阶方程化成一个一阶方程组. 想法是把单个方程

$$d^n\alpha/dt^n = G(t, \alpha, d\alpha/dt, \cdots, d^{n-1}\alpha/dt^{n-1})$$

换成方程组

$$\begin{aligned} d\alpha_1/dt &= \alpha_2, \\ d\alpha_2/dt &= \alpha_3, \\ &\vdots \\ d\alpha_{n-1}/dt &= \alpha_n, \\ d\alpha_n/dt &= G(t, \alpha_1, \cdots, \alpha_{n-1}), \end{aligned}$$

然后认识到这个方程组等价于在一个不同空间上的一个单个的一阶方程. 事实上, 如果我们定义映射 $F = \prec F^1, \cdots, F^n \succ: I \times \boldsymbol{A} \to V = W^n$, 其中 $F^i(t, \boldsymbol{\alpha}) = \alpha_{i+1}, i = 1, \cdots, n-1$, 而 $F^n(t, \boldsymbol{\alpha}) = G(t, \boldsymbol{\alpha})$, 则上面的方程组成为单个方程

$$d\boldsymbol{\alpha}/dt = F(t, \boldsymbol{\alpha}),$$

其中 F 显然对 α 是局部一致利普希茨的. 现在一个函数 $\boldsymbol{f} = \prec f_1, \cdots, f_n \succ: J \to V$ 是此方程的解当且仅当

$$\begin{aligned} f_1' &= f_2, \\ f_2' &= f_3, \\ &\vdots \\ f_{n-1}' &= f_n, \\ f_n' &= G(t, f_1, \cdots, f_n), \end{aligned}$$

那就是说, 当且仅当 f_1 有直到 n 阶的导数, $\psi(f_1) = \boldsymbol{f}$, $f_1^{(n)}(t) = G(t, \psi f_1(t))$. 那个 n 联组的初值条件 $\psi f(t_0) = \boldsymbol{\beta}$ 现在正好是 $\boldsymbol{f}(t_0) = \boldsymbol{\beta}$. 因此对 G 的 n 阶的定理转变成了对 F 的一阶的定理, 故从定理 1.1 和 1.2 得到证明. □

通过 $\prec t_0, \boldsymbol{\beta} \succ$ 的局部解扩张到一个惟一的极大解的结论可以将定理 1.3 用于我们的一阶方程 $d\boldsymbol{\alpha}/dt = F(t, \boldsymbol{\alpha})$ 得到, 并且如果 $G(t, \boldsymbol{\alpha})$ 对 $\boldsymbol{\alpha}$ 是具界 $c(t)$ 的利普希茨连续, 而且 $c(t)$ 连续, 并且 $\boldsymbol{A} = W^n$, 像定理 1.4 那样, 此极大解的定义域为整个 I.

习题

1.1 考虑方程 $d\alpha/dt = F(t, \alpha)$ 的一个特殊情形, 即 $W = \mathbb{R}^2$. 把此方程写成一对包含实(值) 函数和实变量的方程.

1.2 考虑微分方程组

$$dx/dt = t + x^2 + y^2, \quad dy/dt = \cos xy.$$

定义函数 $F : \mathbb{R}^3 \to \mathbb{R}^2$ 使得上面的方程组成为

$$d\alpha/dt = F(t, \alpha),$$

其中 $\alpha = \prec x, y \succ$.

1.3 在上面的习题中证明 F 在 $\mathbb{R} \times A$ 上对 α 是一致利普希茨的, 其中 A 是 $\mathbb{R}^2$ 中的任意有界开集. F 在 $\mathbb{R} \times \mathbb{R}^2$ 上是一致利普希茨的吗 ?

1.4 用解函数 $f = \prec f_1, f_2 \succ$ 表达上面的方程组, 把它写出来. 并写出在证明定理 1.1 中使用的这个方程组的积分形式 .

1.5 在证明定理 1.1 中使用的不动点定理的迭代序列以 f_0 为常值函数 $\overline{\alpha}_0$ 开始, 然后按如下进行 :

$$f_n(t) = \alpha_0 + \int_0^t F(s, f_{n-1}(s))ds.$$

对微分方程

$$dx/dt = t + x \quad [f'(t) = t + f(t)]$$

计算这个序列直到 f_4, 初值条件为 $f(0) = 0$. 这就是说, 取 $f_0 = 0$, 然后按此公式计算 f_1, f_2, f_3 和 f_4. 现在猜一猜解 f 并验证之 .

1.6 对于初值问题

$$dy/dx = x + y^2, y(0) = 0$$

计算迭代 f_0, f_1, f_2 和 f_3. 假设解函数 f 在 0 有一个幂级数展开式, 它的前三个非零项是什么 ?

1.7 在上面习题中对初始条件$f(0) = -1$ 进行计算 .

1.8 对 $f(0) = +1$ 做同样的事 .

1.9 假定 W 为巴拿赫空间, F 和 G 是从 $\mathbb{R}\times W^4$ 到 W 的满足适当的利普希茨条件. 表明二阶方程组

$$\eta''=F(t,\xi,\eta,\xi',\eta'),\quad \xi''=G(t,\xi,\eta,\xi',\eta')$$

如何在我们的标准理论下被变成单个的二阶方程的 .

1.10 把上面的方程转化为一阶方程组, 然后再转化为单个的一阶方程来回答上一个习题 .

1.11 设 θ 为定义在区间 $[0,a]\subset\mathbb{R}$ 上的非负的连续实函数, 并假设存在常数 b 和 $c>0$ 使得

$$\theta(x)\leqslant c\int_0^x\theta(t)dt+bx,\quad \text{其中所有}x\in[0,a]\ .$$

(a) 由归纳法证明, 如果 $m=\|\theta\|_\infty$, 则

$$\theta(x)\leqslant\frac{m(cx)^n}{n!}+\frac{b}{c}\sum_{j=1}^{n}\frac{(cx)^j}{j!},\quad \text{对每个 } n \text{ 成立}\ .$$

(b) 然后证明

$$\theta(x)\leqslant\frac{b}{c}(e^{cx}-1),\quad \text{对所有 } x \text{ 成立}\ .$$

1.12 设 W 为巴拿赫空间, I 为 $\mathbb{R}$ 中一个区间, 并设 F 为从 $I\times W$ 到 W 的一个连续映射. 假设对所有 $t\in I$ 有 $\|F(t,\alpha_0)\|\leqslant b$, 而且对所有 $t\in I$ 和 $\alpha,\beta\in W$ 有

$$\|F(t,\alpha)-F(t,\beta)\|\leqslant c\|\alpha-\beta\|\ .$$

设 f 为经过 $\prec t_0,\alpha_0\succ$ 的整体解, 并令 $\theta(x)=\|f(t_0+x)-\alpha_0\|$. 证明对 $x>0$ 和 $t_0+x\in I$ 有

$$\theta(x)\leqslant\int_0^x\theta(t)dt+bx\ .$$

然后用上一个习题的结果推导出一个比在正文中有的关于 t 离开 t_0 时解 $f(t)$ 的增长的一个更强的界 .

1.13 F 具有上面习题中的假设条件, 证明对通过 $\prec t_0,\alpha_0\succ$ 的解的迭代序列在整个I 收敛, 方程是归纳地证明如果 $f_0=\overline{\alpha}_0$, 且

$$f_n(t)=\alpha_0+\int_0^t F(s,f_{n-1}(s))ds\ ,$$

从而

$$|f_n(t)-f_{n-1}(t)|\leqslant\frac{b}{c}\frac{(ct)^n}{n!}\ .$$

从这些不等式直接证明通过 $\prec t_0,\alpha_0\succ$ 的解满足

$$\|f(t)-\alpha_0\|\leqslant\frac{b}{c}(e^{c|t-t_0|}-1)\ .$$

6.2 对参数的可微依赖性

在某些应用中最重要的是知道方程

$$f'(t) = G(t, f(t)), f(t_1) = \alpha_1$$

的解是如何随初值点 $\prec t_1, \alpha_1 \succ$ 变化的. 为了准确叙述这个问题, 我们固定一个开区间 J, 并像上一节那样令 $\mathrm{u} = B_r(\overline{\alpha}_0) \subset V = \mathfrak{BC}(J, W)$, 并要求 u 中一个通过 $\prec t_1, \alpha_1 \succ$ 的解, 其中 $\prec t_1, \alpha_1 \succ$ 在 $\prec t_0, \alpha_0 \succ$ 的附近. 假定存在一个惟一的解 f, 于是我们有一个映射 $\prec t_1, \alpha_1 \succ \mapsto f$, 正是这个映射的连续性和可微性是我们希望研究的 .

定理 2.1 *设 $L \times U$ 是 $\prec t_0, \alpha_0 \succ$ 在巴拿赫空间 $\mathbb{R} \times W$ 中的一个邻域, 并设 $F(t, \alpha)$ 是从 $L \times U$ 到 W 的一个有界连续映射, 它在 t 上对 α 是一致利普希茨的. 于是存在 $\prec t_0, \alpha_0 \succ$ 的一个邻域 $J \times N$ 具有下列性质: 对每个 $\prec t_1, \alpha_1 \succ \in J \times N$ 存在一个惟一的从 J 到 U 的函数 f, 它是微分方程 $d\alpha/dt = F(t, \alpha)$ 通过 $\prec t_1, \alpha_1 \succ$ 的解, 而且映射 $\prec t_1, \alpha_1 \succ \mapsto f : J \times N \to V$ 为连续 .*

证明 我们简单地重新查验定理 1.1 的计算, 并取 δ 更小一点. 设 $K(t_1, \alpha_1, f)$ 为那个定理的映射但初值点为 $\prec t_1, \alpha_1 \succ$, 故 $g = K(t_1, \alpha_1, f)$ 当且仅当对所有 $t \in J$ 有 $g(t) = \alpha_1 + \int_{t_1}^{t} F(s, f(s))ds$. 显然 K 对每个固定的 f 关于 $\prec t_1, \alpha_1 \succ$ 连续.

如果 N 为球 $B_{r/2}(\alpha_0)$, 则在那个证明中的不等式 (1) 表明 $\|K(t_1, \alpha_1, \overline{\alpha}_0) - \overline{\alpha}_0\| \leqslant \|\alpha_1 - \alpha_0\| + \delta m \leqslant r/2 + \delta m$. 而第二个不等式保持不变. 因此, $f \mapsto K(t_1, \alpha_1, f)$ 是从 u 到 V 的映射, 如果 $\delta c < 1$, 则它是具常数 $C = \delta c$ 的压缩映射. 并且如果 $r/2 + \delta m < (1 - \delta c)r$, 它移动了 u 的中心 $\overline{\alpha}_0$ 一个小于 $(1 - C)r$ 的距离. 这两个对 δ 的新的双重要求等价于

$$\delta < \frac{r}{2(m + cr)},$$

它正好是那个老的值的一半. 以长度为 δ 的 J, 我们现在可以对映射 $K(t_1, \alpha_1, f) : (J \times N) \times \mathrm{u} \to V$ 应用第四章的定理 9.2, 从而定理得证. □

如果我们要求映射 $\prec t_1, \alpha_1 \succ \mapsto f$ 为可微, 由第四章定理 9.4, 只需知道除上面要求外还要求

$$K : (J \times N) \times \mathrm{u} \to V$$

为连续可微就足够了. 要推导出这点, 只需假定 dF 存在并在 $L \times U$ 上一致连续.

定理 2.2 设 $L \times U$ 是 $\prec t_0, \alpha_0 \succ$ 在巴拿赫空间 $\mathbb{R} \times W$ 中的一个邻域, 并设 $F(t, \alpha)$ 为从 $L \times U$ 到 W 的一个有界映射使得 dF 存在, 并有界, 在 $L \times U$ 上一致连续. 于是, 在上面定理的情况下, 解 f 为初值 $\prec t_1, \alpha_1 \succ$ 的连续可微函数.

证明 我们必须证明从 $(J \times N) \times \mathrm{u}$ 到 V 的映射 $K(t_1, \alpha_1, f)$ 为连续可微的, 在此之后像前面评注所说的那样, 我们可以应用第四章的定理 9.4. 现在由 $k(t) = \int_{t_1}^{t} h(s)ds$ 定义的映射 $h \mapsto k$ 是一个从 V 到 V 的有界线性映射, 它显然连续依赖于 t_1, 另外由第三章的定理 14.3, 被积分项映射 $f \mapsto h$ 由 $h(s) = F(s, f(s))$ 定义, 它在 u 上是连续可微的. 将这两个映射复合, 我们看到 $dK^3_{\prec t_1, \alpha_1, f \succ}$ 存在并在 $J \times N \times \mathrm{u}$ 上连续. 现在

$$\Delta K^2_{\prec t_1, \alpha_1, f \succ}(\xi) = \xi ,$$

故 $dK^2 = I$, 而 $\Delta K^1_{\prec t_1, \alpha_1, f \succ}(h) = -\int_{t_1}^{t_1+h} F(s, f(s))ds$, 由此容易得到 $dK^1_{\prec t_1, \alpha_1, f \succ}(h) = -hF(t_1, f(t))$. 这三个偏微分 dK^1, dK^2 和 dK^3 因而存在并在 $J \times N \times \mathrm{u}$ 上连续, 并且由第三章的定理 8.3 得出 $K(t_1, \alpha_1, f)$ 在那里连续可微. □

推论 2.1 如果 s 为 J 中任意一点, 则一个解在 s 的值 $f(s)$ 是其在 t_0 的值的可微函数 .

证明 设 f_α 为通过 $\prec t_0, \alpha \succ$ 的解. 由此定理 $\alpha \mapsto f_\alpha$ 是从 N 到函数空间 $V = \mathfrak{BC}(J, W)$ 的连续可微映射. 但是 $\pi_s : f \mapsto f(s)$ 是个有界线性映射, 因此连续可微是完全平凡的. 复合这两个映射, 我们看出 $\alpha \mapsto f_\alpha(s)$ 在 N 上连续可微. □

把这个解对初值 $\prec t_0, \alpha_0 \succ$ 的连续和可微的依赖性做成整体的形式也是可能的. 下面就是这个定理. 但在这里我们不打算进行证明 .

定理 2.3 设 f 为通过 $\prec t_0, \alpha_0 \succ$ 的极大解, 其定义域为 J, 而设 $[a, b]$ 为 J 的包含 t_0 的任一个有限的闭子区间. 于是存在一个 $\varepsilon > 0$ 使得对任意 $\prec t_1, \alpha_1 \succ \in B_\varepsilon(\prec t_0, \alpha_0 \succ)$, 通过 $\prec t_1, \alpha_1 \succ$ 的整体解的定义域包含了 $[a, b]$, 且这个解在 $[a, b]$ 上的限制是 $\prec t_1, \alpha_1 \succ$ 的连续函数. 如果 F 满足定理 2.2 的假设条件, 则此依赖性是连续可微的 .

最后, 假定 F 连续 (或连续可微地) 依赖于参数 λ, 使得我们在 $M \times I \times A$ 上有 $F(\lambda, t, \alpha)$. 现在对初值问题

$$f'(t) = F(t, f(t)), \quad f(t_1) = \alpha_1$$

的解 f 既依赖于参数 λ 又依赖于初始条件 $f(t_1) = \alpha_1$, 那么如果读者已经充分理解了我们上面的论证, 他将会明白我们能以同样的方式去证明 f 对参数 λ 的依赖也是连续的 (连续可微的). 我们不再进行这些细节讨论了 .

6.3 线性方程

我们现在假定第 1 节的函数 F 是从 $I \times W$ 到 W 的, 并且连续, 对每个固定的 $t, F(t,\alpha)$ 对 α 是线性的. 不难看出我们因此而自动地具有了定理 1.4 的强利普希茨假定条件, 我们将在任何情况下都假定它成立. 在这里这是个关于线性映射的有界性条件: 我们假设 $F(t,\alpha) = T_t(\alpha)$, 其中 $T_t \in \mathrm{Hom}\, W$, 并且 $\|T_t\| \leqslant c(t)$, 对所有 t 成立, 而其中的 $c(t)$ 在 I 上连续 .

像预料的那样, 这种情况下第 1 节的存在性和惟一性理论建立了与一般线性理论的联系. 设 X_0 为所有从 I 到 W 的所有连续函数 $\mathfrak{C}(I, W)$ 的向量空间, 并设 X_1 为所有具有连续的一阶导数的函数的子空间 $\mathfrak{C}^1(I, W)$. 在我们的定理中范数不起作用 .

定理 3.1 如果 $g(t) = f'(t) - F(t, f(t))$, 我们以 $g = sf$ 定义的映射 $S: X_1 \to X_0$ 是个满射的线性映射. 微分方程 $d\alpha/dt = F(t,\alpha)$ 的整体解的集合 N 是 S 的零空间, 因此特别是一个向量空间. 对每个 $t_0 \in I$, 坐标映射 (赋值映射)$\pi_{t_0}: f \mapsto f(t_0)$ 在 N 上的限制是从 N 到 W 的一个同构. 因此 π_{t_0} 的零空间 M 是 N 在 X_1 中的补空间, 故决定了 S 的右逆 R. 映射 $f \mapsto \prec sf, f(t_0) \succ$ 便是从 X_1 到 $X_0 \times W$ 的一个同构, 而这个事实等价于所有前面的论断 .

证明 对 X_0 中任意固定的 g, 令 $G(t,\alpha) = F(t,\alpha) + g(t)$ 并考虑 (非线性) 方程 $d\alpha/dt = G(t,\alpha)$. 由定理 1.3 和 1.4, 它有一个通过任意初始点 $\prec t_0, \alpha_0 \succ$ 的惟一极大解 f, 而且 f 的定义域是整个 I. 那就是说, 对 $X_0 \times W$ 中每个偶对 $\prec g, \alpha \succ$, 存在一个惟一的 $f \in X_1$, 使得 $\prec sf, f(t_0) \succ = \prec g, \alpha \succ$. 映射

$$\prec s, \pi_{t_0} \succ : f \mapsto \prec sf, f(t_0) \succ$$

便是个双射, 并因其明显为线性的, 故它是个同构. 特别 S 是满射的. S 的零空

间 N 是在上述同构下 $\{0\}\times W$ 的逆像; 就是说, $\pi_{t_0}\restriction N$ 是从 N 到 W 的一个同构 .

最后, π_{t_0} 的零空间 M 是 $X_0\times\{0\}$ 在 $\prec S,\pi_{t_0}\succ$ 下的逆像, 而直和分解 $X_1=M\bigoplus N$ 只不过反映了在逆同构下的分解 $X_0\times W=(X_0\times\{0\})\bigoplus(\{0\}\times W)$. 定理证完 . □

对一个给出的 $g\in X_0$ 和一个 $\alpha_0\in W$, 求出一个惟一的 $f\in X_1$ 使得 $S(f)=g$ 和 $f(t_0)=\alpha_0$ 的问题被称作*初值问题*. 在理论的层面上说, 这个问题由上面的定理所解决, 它说那个惟一确定的 f 存在. 从计算的实际层面上说, 这仍是个重要问题 .

$M=M_{t_0}$ 是 N 的补空间这个事实把初值问题破裂为两个独立的子问题. 与 M_{t_0} 相伴的右逆 R 求出 X_1 中的 h 使得 $S(h)=g$ 且 $h(t_0)=0$. 从 N 到 W 的同构 $f\mapsto f(t_0)$ 的逆选出了 X_1 中的 k, 使 $S(k)=0$ 及 $k(t_0)=\alpha_0$. 因此 $f=h+k$. 第一个子问题是 "以齐次初值数据解非齐次方程" 而第二则是 "以非齐次初值数据解齐次方程" 的问题. 在某种意义上说, 初值问题是这两个独立问题的 "直和".

我们现在更仔细地研究齐次方程 $d\alpha/dt=T_t(\alpha)$. 如同我们上面看到的. 它的解空间 N 在每个投影映射 $\pi_t: f\ \mapsto\ f(t)$ 下同构于 W. 设 φ_t 为此同构 (故 $\varphi_t=\pi_t\restriction N$). 现选取某个固定的 $t_0\in I$(我们不妨假定 I 包含 0 并取 $t_0=0$), 并令 $K_t=\varphi_t\circ\varphi_0^{-1}$. 则 $\{K_t\}$ 是 W 自身的线性同构的单参数族, 如果我们令 $f_\beta(t)=K_t(\beta)$, 则 f_β 是 $d\alpha/dt=T_t(\alpha)$ 的通过 $\prec 0,\beta\succ$ 的解. 我们称 K_t 为齐次方程 $d\alpha/dt=T_t(\alpha)$ 的一个*基本解* .

由于 $f'_\beta(t)=T_t(f_\beta(t))$, 我们看出 $d(K_t)/dt=T_t\circ K_t$ 的意义是说此方程对每个 W 中的 β 成立. 然而 $d(K_t)/dt$ 不必是作为在 $\mathrm{Hom}W$ 中的一个范数下极限而存在. 理由是, 我们关于 T_t 的假定并不意味着映射 $t\mapsto T_t$ 从 I 到 $\mathrm{Hom}W$ 是连续的. 如果这个映射是连续, 则映射 $\prec t,A\succ\mapsto T_t\circ A$ 便是从 $I\times\mathrm{Hom}W$ 到 $\mathrm{Hom}W$ 的连续映射, 从而初值问题

$$dA/dt=T_t\circ A,\quad A_0=I$$

在 $\mathfrak{C}^1(I,\mathrm{Hom}W)$ 中有一个惟一的解 A_t. 因为在 β 的赋值是从 $\mathrm{Hom}W$ 到 W 的有界线性映射, 故 $A_t(\beta)$ 是 t 的可微函数, 并且

$$dA_t(\beta)/dt=(dA_t/dt)(\beta)=T_t(A_t(\beta))\ .$$

这表明对所有 β 有 $A_t(\beta)=K_t(\beta)$, 故 $K_t=A_t$. 特别, 基本解 $t\mapsto K_t$ 现在已是

一个到 Hom$W$ 中的可微映射了, 并且 $dK_t/dt = T_t \circ K_t$. 我们已经证明了下面的定理.

定理 3.2 设 $t \mapsto T_t$ 是由 0 的一个区间邻域 I 到 Hom$W$ 的连续映射. 于是微分方程 $d\alpha/dt = T_t(\alpha)$ 的基本解 $t \mapsto K_t$ 是从 I 到 Hom$W$ 的参数弧, 它是初值问题 $dA/dt = T_t \circ A$, $A_0 = I$ 的解.(译注: 作者常将区间 I 与恒同映射 (矩阵)I 在符号上混用. 读者由上下文不难识别而不致混淆.)

利用同构 $K_t = K(t)$ 我们现在可以得到对于非齐次方程 $d\alpha/dt = T_t(\alpha)+g(t)$ 的一个显式表示的解. 我们要求出 f 使得

$$f'(t) - T_t(f(t)) = g(t) .$$

现有 $K'(t) = T_t \circ K(t)$, 故 $T_t = K'(t) \circ K(t)^{-1}$. 从而由第四章习题 8.12 和微分的一般乘积规则 (第三章定理 8.4) 知上面方程的左端恰好是

$$K(t) \left(\frac{d}{dt}[K(t)^{-1}(f(t))] \right) .$$

我们要解的这个方程因而被重写为

$$\frac{d}{dt}[K(t)^{-1}(f(t))] = K(t)^{-1}(g(t)) .$$

于是我们有了一个显然的解, 即使读者觉得我们启发性的论证太技术性了, 他也应该能用微分的方法验证这个解.

定理 3.3 在定理 3.2 的背景下, 函数

$$f(t) = K_t \left[\int_0^t K_s^{-1}(g(s))ds \right]$$

是非齐次初值问题

$$\frac{d\alpha}{dt} = T_t(\alpha) + g(t), \quad f(0) = 0$$

的解. 因此这是 S 的右逆元 R 的一个公式, 它由 S 的零空间 N 的补 M_0 决定.

常系数方程的特殊情形极其重要, 这里的"系数"算子 T_t 是 Hom$W$ 的一个固定 T. 观察到的第一个新事实是当 f 为 $d\alpha/dt = T(\alpha)$ 的一个解时, f' 也是它的一个解. 因为方程 $f'(t) = T(f(t))$ 有一个可微的右端, 微分后得到 $f''(t) = T(f'(t))$. 于是有

引理 3.1 常系数方程$d\alpha/dt = T(\alpha)$ 的解空间在导数算子 D 下不变．

又, 我们从微分方程看出算子 D 在 N 上恰是与 T 的复合. 更准确地说, 方程 $f'(t) = T(f(t))$ 可以重写为 $\pi_t \circ D = T \circ \pi_t$, 并且因为 π_t 在 N 上的限制是 N 到 W 的同构 φ_t, 故从此方程可以解出 T. 因此我们有下面的引理．

引理 3.2 对每个固定的 t, 从 N 到 W 的同构 φ_t 把 N 上的导数算子 D 带到 W 上的算子 T. 即

$$T = \varphi_t \circ D \circ \varphi_t^{-1}$$

现在基本解 K_t 的方程是 $dS/dt = TS$. 在初等微积分中这是个指数函数的方程, 它使我们预期并立刻验证有 $K_t = e^{tT}$(见第四章第 3 节的末尾). $d\alpha/dt = T(\alpha)$ 的通过 $\prec 0, \beta \succ$ 的解因此是函数

$$e^{tT}\beta = \sum_0^\infty t^j \frac{T^j(\beta)}{j!}\ .$$

如果 T 满足一个多项式方程 $p(T) = 0$, 我们知道当 W 为有限维时它必定如此, 于是我们的分析可以大大地向前推进. 假设现在 p 只有实根, 故它的互素因式分解为 $p(x) = \prod_1^k (x - \lambda_i)^{m_i}$. 我们从第一章定理 5.5 知道, W 是直和 $W = \bigoplus_1^k W_i$, 其中 W_i 是变换 $(T - \lambda_i)^{m_i}$ 的零空间, 而每个 W_i 在 T 作用下不变. 这给出了当点 α 在某个零空间 W_i 时的解曲线 $e^{tT}\alpha$ 一个大为简单的形式. 暂且把这样的子空间 W_i 本身就当作 W, 那么我们有 $(T - \lambda I)^m = 0$, 故 $T = \lambda I + R$, 其中 $R^m = 0$, 从而分解 $e^{tT} = e^{t\lambda}e^{tk}$ 连同 e^{tR} 的有限级数展开, 给出了

$$e^{tT}\alpha = e^{t\lambda}\left[\alpha + tR(\alpha) + \cdots + t^{m-1}\frac{R^{m-1}(\alpha)}{(m-1)!}\right].$$

注意, 右端的项数是因子 $(t - \lambda)$ 在多项式 $p(t)$ 中的次数．

在一般情形 $W = \bigoplus_1^k W_i$, 我们有 $\alpha = \sum_1^k \alpha_i$, $e^{tT}(\alpha) = \sum_1^k e^{tT}(\alpha_i)$, 而每个 $e^{tT}(\alpha_i)$ 具有上面的形式. $f'(t) = T(f(t))$ 的经过一般点 $\prec 0, \alpha \succ$ 的解因而是形如 $t^j e^{t\lambda_i}\beta_{ij}$ 的项的有限和, 项的个数等于多项式 p 的次数．

如果 W 是个复巴拿赫空间, 那么关于 p 只有实根的限制是多余的. 我们除了 λ 为复数值外得到了完全一样的公式. 这在解曲线的行为上引进了更多的变化, 这是因为外部的指数因子 $e^{t\lambda} = e^{t\mu}e^{it\nu}$ 当 $\nu \neq 0$ 时有了一个周期的因子．

综合起来我们已经证明了下面的定理.

定理 3.4 如果 W 是个实或复的巴拿赫空间, $T \in \mathrm{Hom}W$, 于是初值问题 $f'(t) = T(f(t)), f(0) = \beta$ 在 W 中的解曲线是

$$f(t) = e^{tT}\beta = \sum_0^\infty \frac{t^j}{j!}T^j(\beta) .$$

如果 T 满足多项式方程 $(T-\lambda)^m = 0$, 则

$$f(t) = e^{t\lambda}\left[\beta + tR(\beta) + \cdots + \frac{t^{m-1}}{(m-1)!}R^{m-1}(\beta)\right] ,$$

其中 $R = T - \lambda I$. 如果 T 满足多项式方程 $p(T) = 0$, p 具有互素因式分解 $p(x) = \prod_1^k (x-\lambda_i)^{m_i}$, 则 $f(t)$ 是 k 个上面形式的项的和, 故有形式

$$f(t) = \sum_{i,j} t^j e^{t\lambda_i}\beta_{ij} ,$$

其中右端的项数是多项式 p 的次数, 每个 β_{ij} 是固定的 (常值) 向量.

留意当 $t \to +\infty$ 时 $f(t)$ 的渐近行为是怎样被多项式的根 λ_i 控制的是一件重要的事. 我们首先把我们限制于通过某个子空间 W_i 的一个向量 α 的解上, 它等于假设了 $(T-\lambda)^m = 0$. 于是, 如果 λ 有个正的实部, 故 $e^{t\lambda} = e^{t\mu}e^{it\nu}, \mu > 0$, 则 $\|f(t)\|$ 按指数样子趋向无穷大. 如果 λ 有负的实部, 则当 $t \to \infty$ 时 $f(t)$ 趋向于零 (但其范数当 $t \to -\infty$ 时成为指数式快速的无穷大). 如果 λ 的实部为 0. 则如果 $m > 1, \|f(t)\|$ 像 t^{m-1} 那样趋向无穷大. 因此 f 在整个 $\mathbb{R}$ 上有界的惟一方式是 λ 的实部为 0 且 $m = 1$, 这时 f 是周期的. 相似地, 在一般情形, 这时 $p(T) = \prod_1^k (T-\lambda_n)^{m_n} = 0$, 那相似的结论仍是正确的, 即所有的解曲线在整个 $\mathbb{R}$ 上有界当且仅当根 λ_n 全都是纯虚数且所有的重数 m_n 为 1.

习题

3.1 设 I 为 $\mathbb{R}$ 中一个开区间, W 为赋范线性空间. 设 $F(t,\alpha)$ 为从 $I \times W$ 到 W 的一个连续函数, 对每个固定的 t 它对 α 是线性的. 证明存在一个函数 $c(t)$, 它在每个包含在 I 中的闭区间 $[a,b]$ 是有界的, 并使得对所有的 α 和 t 有 $\|F(t,\alpha)\| \leqslant c(t)\|\alpha\|$. 那么证明可以使得 c 为连续.(你可能要用到海涅 - 波雷尔性质:如果 $[a,b]$ 被一组开区间覆盖, 则某个有限的子组已经覆盖了 $[a,b]$.)

3.2 在正文中我们略去了检验 $f \mapsto f^{(n)} - G(t,f,f',\cdots,f^{(n-1)})$ 是从 X_n 到 X_0 的满射. 通过在简约到一阶方程组的过程中追寻满射性的方法证明上述结论.

3.3 假设在算子

$$Tf = \sum_0^n a_i f^{(i)}$$

中系数 $a_i(t)$ 全都在 $\mathfrak{C}^1$ 中. 证明 T 的零空间 N 是 $\mathfrak{C}^{n+1}$ 中的子空间. 叙述这个定理的一个推广并粗略地说明为什么它是对的.

3.4 假设 W 为巴拿赫空间, $T \in \mathrm{Hom}W$, β 是 T 的特征向量, 其特征值为 r. 证明常系数方程 $d\alpha/dt = T(\alpha)$ 通过 $\prec 0, \beta \succ$ 的解是 $f(t) = e^{tr}\beta$.

3.5 假定 W 还是有限维的, 并有一个基 $\{\beta_i\}_1^n$, 由 T 的特征向量组成, 其相应的特征值为 r_i. 用 α 的基展开式求通过 $\prec 0, \alpha \succ$ 的解公式 .

3.6 线性方程 $d\alpha/dt = T_t(\alpha)$ 的一个非常重要的特殊情形是算子函数 T_t 为周期的. 例如, 假设对所有 t 有 $T_{t+1} = T_t$. 证明这时对所有 t 和 n 成立 $K_{t+\eta} = K_t(K_1)^n$.

其次我们假定 K_1 具有一个对数, 故可写成 $K_1 = e^A$, A 为 $\mathrm{Hom}W$ 中某个元.(我们从第四章习题 11.19 知道, 如果 W 为有限维这总是可能的.) 证明现在可将 K_t 写为形式

$$K_t = B(t)e^{tA} ,$$

其中 $B(t)$ 是以周期为 1 的周期映射 .

3.7 继续进行上面的习题, 假定现在 W 为有限维复向量空间. 利用正文中给出的对 $e^{tA}\beta$ 的分析, 证明微分方程 $d\alpha/dt = T_t(\alpha)$ 有一个周期解 (具任何一种周期) 仅当 K_1 有一个绝对值为 1 的特征值. 又证明如果 K_1 有一个 n 次单位根的特征值, 则此微分方程有一个周期为 n 的周期解.

3.8 写出定理 3.3 中在常系数情形下公式的特殊形式 .

3.9 以第一章的定理 5.3 的观点观察定理 3.1 中是事实是颇有意思的. 假定 $S : X_1 \to X_0$ 为满射, 且其零空间 N 在坐标 (赋值) 映射 π_{t_0} 下同构于 W. 证明如果 M 是 π_{t_0} 在 X_1 中的零空间, 则 $S \restriction M$ 是到 X_0 上的一个同构, 这可由应用此定理而得证 .

6.4 n 阶线性方程

n 阶线性微分方程是方程

$$d^n\alpha/dt^n = G(t, \alpha, d\alpha/dt, \cdots, d^{n-1}\alpha/dt^{n-1}) ,$$

其中 $G(t, \boldsymbol{\alpha}) = G(t, \alpha_1, \cdots, \alpha_n)$ 现在是从 $V = W^n$ 到 W 在每个 $t \in I$ 为线性的. 我们把它转化为一阶方程 $d\boldsymbol{\alpha}/dt = F(t, \boldsymbol{\alpha})$ 就像上面做过的那样, 这里的 F 是从 $I \times V$ 到 V 的对第二个变量 $\boldsymbol{\alpha}$ 为线性的映射, $F(t, \boldsymbol{\alpha}) = T_t(\boldsymbol{\alpha})$.

我们对定理 1.5 的证明表明 $\mathfrak{C}^n(I,W)$ 中的函数 f 是 n 阶方程 $d^n\alpha/dt^n = G(t,\alpha,\cdots,d^{n-1}\alpha/dt^{n-1})$ 的一个解当且仅当 n 联组 $\psi f = \prec f,f',\cdots,f^{(n-1)}\succ$ 是一阶方程 $d\boldsymbol{\alpha}/dt = F(t,\boldsymbol{\alpha}) = T_t(\boldsymbol{\alpha})$ 的解. 我们知道后面的这些解形成了 $\mathfrak{C}^1(I,W^n)$ 的一个向量子空间 $\boldsymbol{N}$, 并且因为映射 $\psi: f \mapsto \prec f,f',\cdots,f^{(n-1)}\succ$ 是从 $\mathfrak{C}^n(I,W)$ 到 $C^1(I,W)$ 线性的, 从而此 n 阶方程的解集合 N 是 $\mathfrak{C}^n(I,W)$ 的子空间, 而 $\psi\restriction N$ 是从 N 到 $\boldsymbol{N}$ 的同构. 由于坐标赋值映射 $\varphi_t = \pi_t\restriction\boldsymbol{N}$ 对每个 t 是从 $\boldsymbol{N}$ 到 W^n 的同构 (定理 3.1), 由此得到映射

$$\pi_t\circ\psi: f \ \mapsto \prec f(t),f'(t),\cdots,f^{(n-1)}(t)\succ$$

将 N 同构地带到 W^n. 其在 $\mathfrak{C}^n$ 中零空间 M_t 是 N 的补, 这与以前的情况一样. 这里的 M_t 是 $\mathfrak{C}^n(I,W)$ 中使 $f(t)=\cdots=f^{(n-1)}(t)=0$ 的函数 f 的集合 .

我们现在考虑特殊情况 $W=\mathbb{R}$. 对每个固定的 t,G 是从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的线性映射, 即 $(\mathbb{R}^n)^*$ 中的一个元, 而它关于标准基的坐标集合是一个 n 联组 $\boldsymbol{k} = \prec k_1,\cdots,k_n\succ$. 由于此线性映射随 t 连续变化, 故这个 n 联组 $\boldsymbol{k}$ 随 t 而连续变化. 因此当我们考虑到 t 时我们有在 I 上连续实函数的 n 联组 $k(t) = \prec k_1(t),\cdots,k_n(t)\succ$ 使得

$$G(t,x_1,\cdots,x_n)=\sum_{i=1}^n k_i(t)x_i\ .$$

n 阶微分方程

$$d^n\alpha/dt^n = G(t,\alpha,\cdots,d^{n-1}\alpha/dt)$$

的解空间 N 恰是线性变换 $L:\mathfrak{C}^n(I,\mathbb{R})\to\mathfrak{C}^0(I,\mathbb{R})$ 的零空间, 这里的 L 被定义为

$$(Lf)(t)=f^{(n)}(t)-k_n(t)f^{(n-1)}(t)-\cdots-k_1(t)f(t)\ .$$

如果我们将指标移位使其与导数的阶一致, 并且如果我们让 $f^{(n)}$ 也有一个系数函数, 则我们的 n 阶线性微分算子 L 的样子是

$$(Lf)(t)=a_n(t)f^{(n)}(t)+\cdots+a_0(t)f(t)\ .$$

只要假设 $a_n(t)$ 恒不为零则给出 $f^{(n)}$ 一个系数函数 a_n 并没有改变什么, 这是因为可以用它做除法又给出我们原来所研究的形式. 称此为正规情形. 在 $a_n(t)$ 对某个 t 为零时的奇异情形需要做进一步的研究, 我们将不再深入下去.

概括在这种情形下一般线性理论告诉我们的东西 .

定理 4.1 L 是从空间 $\mathfrak{C}^n(I)$ 到 $\mathfrak{C}^0(I)=\mathfrak{C}(I)$ 的满射的线性变换, 其中 $\mathfrak{C}^n(I)$ 是 I 上有直到 n 阶连续导数的实函数全体, $\mathfrak{C}(I)$ 是 I 上连续函数的全体构成的空间. 它的零空间 N 是我们原来的微分方程的解空间. 对每个 $t_0\in I$, 映射 $\varphi_{t_0}\circ\psi: f\mapsto\prec f(t_0),\cdots,f^{(n-1)}(t_0)\succ$ 在 N 上的限制是从 N 到 $\mathbb{R}^n$ 的一个同构, 而 $\mathfrak{C}^n$ 中使得 $f(t_0)=\cdots=f^{(n-1)}(t_0)=0$ 的函数 f 的集合 M_{t_0} 因而是 N 在 $\mathfrak{C}^n(I)$ 中的补空间, 并决定了 L 的一个线性右逆 .

对给出的 g 实际 "解" 出微分方程 $L(f)=g$ 的 f 的问题分成了两个部分. 首先我们必须求出 L 的零空间 N, 即必须解齐次方程 $L(f)=0$. 因为 N 是个 n 维向量空间, 所以进行刻画解的问题等于找出一组基, 这显然是进行解的最有效的办法. 我们的第一个问题因而是找出 $L(f)=0$ 的 n 个线性无关的解 $\{u_i\}_1^n$. 我们的第二个问题是找出 L 的一个右逆, 就是说, 选择一个f 的线性方法使得 $L(f)=g$ 对每个 g 成立. 这里显然要做的事就是力图使定理 3.3 的公式能给出一个实际的计算. 如果 v 是 $L(f)=g$ 的一个解, 则当然地所有解的集合便是仿射子空间 $N+v$.

我们从第一个问题开始着手, 即找出 $L(f)=0$ 的解的一个基 $\{u_i\}_1^n$. 可惜并没有现成的一般方法, 那么我们不得不满足于部分的成功了. 我们将看到我们能容易地直接解出一阶方程, 并且如果我们能找出 n 阶方程的一个解则可把此问题简约为解一个 $n-1$ 阶的方程. 另外, 在具常系数的算子 L 的非常重要的特殊情形中, 定理 3.4 给出了一个完全的显式解 .

一阶齐次线性方程可以被写成形式 $y'+a(t)y=0$, 其中 y' 的系数已用除法去掉. 除以 y 并记住 $y'/y=(\log y)'$, 我们看出, 至少是形式地看出由 $\log y=-\int a(t)dt$ 或 $y=e^{-\int a(t)dt}$ 给出的一个解, 我们可通过检验来确定它. 那么, 方程 $y'+y/t=0$ 便有一个解 $y=e^{-\log t}=1/t$, 读者或许已经知道它了 .

假设现在 L 为 n 阶算子并且我们已知道 $L(f)=0$ 的一个解 u. 我们的问题便是求出 $n-1$ 个解 $v_1,\cdots,v_{n-1}$, 它们相互无关也与 u 无关. 有理由猜测这将被确定为一个 $n-1$ 阶方程的解. 我们试图找出一个形如 $c(t)u(t)$ 的第二个解 $v(t)$, 其中 $c(t)$ 是个未知函数. 我们的动机部分在于这样的一个解将自动地与 u 无关, 除非 $c(t)$ 最后成了一个常数 .

如果 $v(t)=c(t)u(t)$, 则 $v'=cu'+c'u$, 一般地

$$v^{(j)}=\sum_{i=0}^{j}\binom{j}{i}c^{(i)}u^{(j-i)}\ .$$

如果我们写出 $L(v)=\sum_0^n a_j(t)v^{(j)}(t)$, 并对包含 $c(t)$ 的这些项并项, 得到

$$L(v)=c(t)\sum_0^n a_j u^{(j)}+\text{包含 } c',\cdots,c^{(n)}\text{的那些项}$$
$$=cL(u)+S(c')=S(c') ,$$

其中 S 是一个确定的 $n-1$ 阶线性微分算子, 它可由上面公式显式地算出来. 我们断言解 $S(f)=0$ 就是解原来的方程. 因为假设 $\{g_i\}_1^{n-1}$ 是 S 的零空间的一个基, 令 $c_i(t)=\int_0^t g_i$. 于是 $L(c_i u)=S(c_i')=S(g_i)=0, i=1,\cdots,n-1$. 又 $u,c_1u,\cdots,c_{n-1}u$ 是无关的, 这是因为如果 $u=\sum_1^{n-1}k_ic_iu$, 则 $1=\sum_1^{n-1}k_ic_i(t)$, 从而 $0=\sum_1^{n-1}k_ic_i'(t)=\sum_1^{n-1}k_ig_i(t)$, 这与集合 $\{g_i\}$ 的无关性相矛盾.

于是我们已经证明如果我们能找到 n 阶方程 $Lf=0$ 的一个解 u, 则其全解被简化为 $n-1$ 阶方程 $Sf=0$(虽然我们独立性论证是粗略一点).

这个约化程序没有结合一阶方程的解从而建立 n 阶方程的一组无关的解, 这是因为粗略地说, 它 "从上往下做而不是从下往上". 为了能使这样的结合成功, 我们不得不从给出的 n 阶算子找到一个一阶算子 S, 使得 $N(S)\subset N(L)$, 我们一般做不到这点. 然而当 L 中的系数函数全为常数时我们能做到, 但是事实上我们将按不同的方法进行.

同时值得注意的是对一个二阶方程 $Lf=0$, 如果我们能求出一个解来则可以完全解出它, 原因是上述论证把剩下的问题约化成了一个一阶方程, 而这由我们前面看到的, 可以由一个积分把它解出. 例如, 考虑方程 $y''-2y/t^2=0$ 在任意不含 0 的区间 I 上的解, 有 $a_0(t)=1/t^2$ 在 I 上连续. 由观察可知 $u(t)=t^2$ 是个解. 于是我们知道我们能够找到一个形如 $v(t)=t^2c(t)$ 的解, 它与 $u(t)$ 无关, 而问题化作一个对 c' 的一阶方程. 事实上我们有 $v'=t^2c'+2tc$ 和 $v''=t^2c''+4tc'+2c$, 故 $L(v)=v''-2v/t^2=t^2c''+4tc'$, 而 $L(v)=0$ 当且仅当 $(c')'+(4/t)c'=0$. 因此

$$c'=e^{-\int 4dt/t}=e^{-4\log\, t}=1/t^4, c=1/t^3$$

(差一个纯量因子; 我们只需要一个基!), 而 $v=t^2c(t)=1/t$. (读者或愿验证一下这是个所期望的解.) 算子 $L(f)=f''-2f/t^2$ 的零空间因而由 $\{t^2,1/t\}$ 线性张成.

我们现在转向一个重要而易于处理的情形, 即微分算子

$$Lf=a_nf^{(n)}+a_{n-1}f^{(n-1)}+\cdots+a_0f ,$$

其中系数 a_i 为常数, a_n 则不妨取为 1. 使得这种情形易于处理在于现在 L 是导数算子 D 的多项式. 那就是说, 如果 $Df = f'$, 故 $D^j f = f^{(j)}$, 于是 $L = p(D)$, 其中 $p(x) = \sum_0^n a_i x^i$.

处理这个方程最优美但并非最初等的方法是转移到 $\mathbb{R}^n$ 上的等价一阶方程组 $d\boldsymbol{x}/dt = T(\boldsymbol{x})$, 故可用上一节的合适的理论 .

定理 4.2 如果 $p(t) = (t-b)^n$, 则常系数 n 阶方程 $p(D)f = 0$ 的解空间 N 具有基

$$\{e^{bt}, te^{bt}, \cdots, t^{n-1}e^{bt}\} .$$

如果 $p(t)$ 是一个多项式有互素因式分解 $p(t) = \prod_1^k p_i(t)$, 其中每个 $p_i(t)$ 具有上面的形式, 则常系数方程 $p(D)f = 0$ 解空间具有基 $\cup B_i$, 其中 B_i 是对方程 $p_i(D)f = 0$ 的解空间 N_i 的由上面写出的基 .

证明 我们知道映射 $\psi : f \mapsto \prec f, f', \cdots, f^{n-1} \succ$ 是从 $p(D)$ 的零空间 N 到 $d\boldsymbol{x}/dt - T(\boldsymbol{x})$ 的零空间 $\boldsymbol{N}$ 的一个同构. ψ 显见是与微分交换的, $\psi(Df) = \prec f', \cdots, f^{(n)} \succ = D\psi(f)$, 并且由引理 3.1 我们也知道 $\boldsymbol{N}$ 在 D 下不变, 那么便得到 (也可容易直接验证)N 在 D 下不变. 由引理 3.2 我们有 $T = \varphi_t \circ D \circ \varphi_t^{-1}$, 它不过是在说同构 $\varphi_i : \boldsymbol{N} \to \mathbb{R}^n$ 将 $\boldsymbol{N}$ 上的算子 D 变成了 $\mathbb{R}^n$ 上的算子 T. 总合起来, $\varphi_t \circ \psi$ 将 N 上的 D 变成了 $\mathbb{R}^n$ 上的 T, 并且由于在 N 上 $p(D) = 0$, 由此得出在 $\mathbb{R}$ 上 $p(T) = 0$.

在定理 3.4 中我们看到如果 $p(T) = 0$ 和 $p = (t-b)^n$, 于是 $d\boldsymbol{x}/dt = T(\boldsymbol{x})$ 的解空间 $\boldsymbol{N}$ 由形如

$$e^{bt}\boldsymbol{x}, \cdots, t^{n-1}e^{bt}\boldsymbol{x}$$

的向量张成. 这个 n 联组的函数 $\boldsymbol{g}$ 的第一个坐标在 $\boldsymbol{N}$ 中构成了空间 N(在 $f = \psi^{-1}\boldsymbol{g}$ 的同构下), 因此我们看出了 N 是由函数 $e^{bt}, \cdots, t^{n-1}e^{bt}$ 张成的. 由于 N 是 n 维的并且这里有 n 个这些函数, 故这个去进行张成 N 的集合构成一个基.

定理的剩余部分可以被看作为上面已证部分和第一章定理 5.5 对 N 上方程 $p(D) = 0$ 的直接应用的结合, 或者作为在 ψ^{-1} 下把上节中已经对 $\boldsymbol{N}$ 建立的事实转移到了 N . □

如果多项式 p 的根不全都是实的, 我们则不得不求助于在第四章 11 节的习题阐述的复化理论. 除去最后的一个步骤外, 结果都是一样的. 一个必须应用的额外事实是一个实算子作用于实向量空间 Y 的零空间恰好是作用在 Y 的复化 $Z = Y \bigoplus iY$ 上的 T 的复化 S 的零空间与 Y 的交. 这意味着如果 $p(t)$ 是个实

系数的多项式, 则我们从复解的实部得出了 $p(D)f=0$ 的实解. 为了确切地看出其中的意思, 假设 $q(x)=(x^2-2bx+c)^m$ 是 $p(x)$ 中在 $\mathbb{R}$ 上的一个互素因子, 其中 $x^2-2bx+c$ 在 $\mathbb{R}$ 不可约. 在 $\mathbb{C}$ 上, $q(x)$ 分解为 $(x-\lambda)^m.\ (x-\overline{\lambda})^m$, 其中 $\lambda=b+i\omega$, $\omega^2=c-b^2$. 由我们上面的一般理论得到 $q(D)$ 的复 $2m$ 维零空间是由

$$\{e^{\lambda t},te^{\lambda t},\cdots,t^{m-1}e^{\lambda t},e^{\overline{\lambda}t},te^{\overline{\lambda}t},\cdots,t^{m-1}e^{\overline{\lambda}t}\}\ .$$

复张成的.

这 $2m$ 个函数的复线性组合的实部是一个由上面这些函数的实部和以 i 乘这些函数后的实部张成的 $2m$ 维实向量空间. 就是说实算子 $q(D)$ 的零空间是由

$$\{e^{bt}\cos\omega t,te^{bt}\cos\omega t,\cdots,t^{m-1}e^{bt}\cos\omega t;e^{bt}\sin\omega t,\cdots,t^{m-1}e^{bt}\sin\omega t\}\ .$$

张成的 $2m$ 维实空间. 由于有 $2m$ 个这种函数, 它们必为无关, 必形成 $q(D)f=0$ 的实解空间的一组基. 因此有

定理 4.3 如果 $p(t)=(t^2+2bt+c)^m, b^2<c$, 则 $2m$ 阶常系数方程 $p(D)f=0$ 的解空间有基

$$\{t^ie^{bt}\cos\omega t\}_{i=0}^{m-1}\bigcup\{t^ie^{bt}\sin\omega t\}_{i=0}^{m-1}\ ,$$

其中 $\omega^2=c-b^2$. 对任意的实系数多项式 $p(t)$, 如果 $p(t)=\prod_1^k p_i(t)$ 为它的互素因子分解, 分解为线性因子的幂和不可约二次因子的幂, 于是 $p(D)f=0$ 的解空间 N 具有基 $\bigcup_1^k B_i$, 其中 B_i 是 $p_i(0)$ 的零空间的基; 如果 $p_i(t)$ 是一个不可约二次多项式的幂则 B_i 是我们上面展示的那样; 如果 p_i 是线性因子的幂, 则 B_i 是定理 4.2 中的基 .

例如, 假如我们要求 $D^4-1=0$ 的零空间的基. 这里的 $p(x)=x^4-1=(x-1)(x+1)(x-i)(x+i)$, 复解空间的基便是 $\{e^t,e^{-t},e^{it},e^{-it}\}$. 因为 $e^{it}=\cos t+i\sin t$, 故其实解空间的基是 $\{e^t,e^{-t},\cos t,\sin t\}$.

对 $D^3-1=0$ 的同一问题给出了

$$\begin{aligned}p(x)&=x^3-1=(x-1)(x^2+x+1)\\&=(x-1)\left(x+\frac{1+i\sqrt{3}}{2}\right)\ \left(x+\frac{1-i\sqrt{3}}{2}\right)\ ,\end{aligned}$$

故此复解空间的基是

$$\{e^t,e^{-[(1+i\sqrt{3})/2]t},\quad e^{-[(1-i\sqrt{3})/2]t}\},$$

而此实解空间的基是

$$\{e^t, e^{-t/2}\cos(\sqrt{3}t/2),\quad e^{-t/2}\sin(\sqrt{3}t/2)\}\ .$$

* 上面的结果提示我们, 所有常系数齐次线性微分方程的所有实解的集 $\mathfrak{a}$ 包含了 t^i, e^{rt}, $\cos\omega t, \sin\omega t$, 其中 i, r, ω 任意, 并且对加法和乘法是封闭的, 实际上是由这些函数生成的一个代数 .

我们可以容易证明这个猜测. 首先考虑和. 假设 $T(f)=0$ 和 $S(g)=0$, 其中 T 和 S 为两个这种常系数算子. 则因为 S 和 T 可交换, 则 $f+g$ 在 $S\circ T$ 的零空间中 :

$$(S\circ T)(f+g)=(S\circ T)(f)+(S\circ T)(g)=S(Tf)+T(Sg)=0+0=0\ .$$

我们知道 S 和 T 之所以可交换是因为它们都是 D 的多项式 .

为了处理乘积, 我们首先必须认识到所有三角函数 $\sin at, \cos at$ 的线性张成空间是个代数. 换句话说, 这些函数的有限积是这些函数的线性组合. 这是某一类三角恒等式起的作用. 诸如 $2\sin x\cos x=\sin(x+y)+\sin(x-y)$, 读者无疑不得不努力对付它们. (当我们被允许用复指数处理它们时则所有秘密之处都已消失.) 于是我们观察到代数 $\mathfrak{a}$ 中任意的函数都是对某些 i, r, ω 的函数形式 $t^ie^{rt}\sin\omega t$ 或 $t^ie^{rt}\sin\omega t$ 作为项的有限和. 我们可以列出有这样一个函数在其零空间的算子 T, 而我们这些项的有限和则在这些算子 T 的复合算子的零空间中, 这是我们第一部分论证中得到的 .

我们想多说几句. 函数 $t^i, e^{rt}, \sin\omega t, \cos\omega t$ 以及它们乘积的和可以被证明正好是那些连续函数 $f:\mathbb{R}\to\mathbb{R}$ 使得 f 的所有平移的集合有一个有限维的张成. 那就是说, 如果我们定义通过 x 的平移 K_x 为 $(K_xf)(t)=f(t-x)$, 于是恰好是上面的函数 f 我们有 $\{K_xf: x\in\mathbb{R}\}$ 的线性张成是有限维的. 对完全相同的函数类的第二种特征描述不会是偶然所得. 其部分秘密所在是因为常系数算子 T 正好是那些与平移可交换的线性微分算子. 就是说, 如果 T 是个线性微分算子, 则对所有 x 有 $T\circ K_x=K_x\circ T$ 当且仅当 T 具有常系数. 我们在前一章已经注意到, 如果 $T\circ S=S\circ T$, 则 T 的零空间是在 S 下不变的. 因此常系数算子 T 的零空间在所有平移下不变: 对所有 x, $K_x[N]\subset N$. 现在我们从微分方程理论知道 N 是有限维的. 因此 N 中的函数是那些使它们的平移张成了有限维空间的函数 !

找出与算子 T 交换的 S 从而使线性算子 T 的零空间 N 为 S - 不变, 这种获取关于 N 的更多的信息的手段在高等数学中用得很多. 具有特殊重要性的是

我们有一组可交换算子 S 的情形, 这就像我们前面的算子 $S = K_x$ 的那种情形.

我们还没有证明的是, 如果一个连续函数 f 的平移生成了有限维向量空间, 则 f 是某个常系数算子 $p(D)$ 的零空间中的元. 这是美妙的, 它有赖于证明如果 $\{K_t\}$ 是在一有限维空间上的线性变换的单参数族满足 $K_{s+t} = K_s \circ K_t$ 及当 $t \to 0$ 时 $K_t \to I$, 则存在一个 $S \in \mathrm{Hom}V$ 使得 $K_t = e^{tS}$. *

习题

求下列方程的解 .

4.1 $x'' - 3x' + 2x = 0$.

4.2 $x'' + 2x' - 3x = 0$.

4.3 $x'' + 2x' + 3x = 0$.

4.4 $x'' + 2x' + 1 = 0$.

4.5 $x''' - 3x'' + 3x' - x = 0$.

4.6 $x''' - x = 0$.

4.7 $x^{(6)} - x'' = 0$.

4.8 $x''' = 0$.

4.9 $x''' - x'' = 0$.

4.10 解初值问题 $x'' + 4x' - 5x = 0, x(0) = 1, x'(0) = 2$.

4.11 解初值问题 $x''' + x' = 0, x(0) = 0, x'(0) = -1, x''(0) = 1$.

4.12 求方程 $4t^2x'' + x = 0$ 的一个解 u : 试一试 $u(t) = t^n$, 然后令 $v(t) = c(t)u(t)$, 像在正文中那样求第二个解 .

4.13 解 $t^3x''' - 3tx' + 3x = 0$, 试一试 $u(t) = t^n$.

4.14 解 $tx'' + x' = 0$.

4.15 解 $t(x''' + x') + 2(x'' + x) = 0$.

4.16 已知 $e^{-bt}\cos \omega t$ 和 $e^{-bt}\sin \omega t$ 是一个二阶线性微分方程的解, 也知道它们在 0 的值分别为 1 和 0; 我们还知道它们是线性无关的. 为什么?

4.17 求一个常系数微分方程使得下列函数为它的解:$t^2, \sin t, t^2\sin t$.

4.18 如果 f 和 g 是二阶线性微分方程 $u'' + a_1u' + a_2u = 0$ 的无关解, 其中方程的系数为连续函数, 于是我们知道向量 $\prec f(x), f'(x) \succ$ 与 $\prec g(x), g'(x) \succ$ 在每个 x 是无关的. 反过来证明, 如果两个函数具前面的性质, 则它们是一个二阶微分方程的解 .

4.19 用在正文中讨论的阶的约化程序解方程 $(D-a)^3f=0$; 从一个显然的解 e^{at} 开始 .

6.5 解非齐次方程

我们现在到了解非齐次方程$L(f)=g$ 问题的时候. 我们将简短地描述一个实用的方法, 它在某些时候简易可行, 我们还要描述一个理论上的方法, 它永远可行, 但却难于应用. 后一种只是把定理 3.3 翻译成矩阵的语言 .

我们首先考虑常系数方程 $L(f)=g$ 的一个特殊情形, 这时 g 本身是在某个常系数算子 S 的零空间中. 简单的例子是 $y'-ay=e^{bt}$(或 $y'-ay=\sin bt$), 其中 $g(t)=e^{bt}$ 是在 $S=(D-b)$ 的零空间中. 在这种情形中, 解 f 必在 $S\circ L$ 的零空间中, 这是因为 $S\circ L(f)=s(g)=0$. 我们知道了所有这些函数是什么样子后, 我们的问题便是在它们中选出 f 使得 $L(f)$ 等于给出的 g .

暂且假定多项式 L 和 S(D 的多项式) 没有公因式. 于是我们由此知道 L 在 S 的零空间 N_S 上是一个同构, 因而存在一个 $f\in N_S$ 使得 $Lf=g$. 由于我们有 N_S 的一组基, 我们可以对 L 在 N_S 上的作用构造一个矩阵, 从而由解矩阵方程求出 f, 然而最简单的做法是取此基的一个以未知量为系数的一般线性组合, 让 L 作用于它, 然后看出系数必须怎样才能给出 g .

例如要解 $y'-ay=e^{bt}$, 我们可试用 $f(t)=ce^{bt}$, 并应用

$$L:(D-a)(ce^{bt})=(b-a)ce^{bt}\overset{?}{=}e^{bt}\ ,$$

于是看出 $c=1/(b-a)$.

另外, 解 $y'-ay=\cos bt$, 我们看出 $\cos bt$ 在 $S=D^2+b^2$ 的零空间中, 而此零空间有基 $\{\sin bt,\cos bt\}$. 因此令 $f(t)=c_1\sin bt+c_2\cos bt$, 然后解 $(D-a)f=\cos bt$, 得到

$$(-ac_1-bc_2)\sin bt+(bc_1-ac_2)\cos bt=\cos bt\ ,$$

$$-ac_1-bc_2=0,\quad bc_1-ac_2=1\ ,$$

从而

$$f(t)=\frac{b}{a^2+b^2}\sin bt-\frac{a}{a^2+b^2}\cos bt\ .$$

当 L 和 S 具有公因式时, 这种情况就更复杂, 但相似的程序可以作出证明. 现在必须引进一个附加因式 t^i, 其中 i 为 L 中公因式出现的数. 例如, 解

$(D-r)^2f=e^{rt}$, 我们有 $S\circ L=(D-r)^3$, 因此我们必须令 $f(t)=ct^2e^{rt}$. 于是我们的方程变为

$$(D-r)^2ct^2e^{rt}=2ce^{rt}\stackrel{?}{=}e^{rt}\ ,$$

因而 $c=\frac{1}{2}$.

对于 $(D^2+1)f=\sin t$ 我们必须设 $f(t)=t(c_1\sin t+c_2\cos t)$, 而后我们将其算出找到 $c_1=0,\ c_2=-\frac{1}{2}$, 使得 $f=-\frac{1}{2}t\cos t$.

这种程序自然地被叫做*待定系数法*, 它违背了我们关于求解过程是一个求线性右逆这个原理. 的确, 它并非可用于出现在右端的任意 g 的单一过程, 而是随算子 S 而改变. 然而当它能用时, 它是最简单的计算显式解的办法 .

我们下面要描述一般理论上的一个方法, 它被称作*参数变值法*, 是一个 L 的右逆, 而且确实能用于每一个 g. 另外, 它回到了一般的 (变系数) 线性 n 阶算子 L :

$$(Lf)(t)=\sum_0^n a_i(t)f^{(i)}(t)\ .$$

我们假定已知 L 的零空间 $\boldsymbol{N}$; 就是说, 假定我们知道了齐次方程 $Lf=0$ 的 n 个线性无关的解 $\{u_i\}_1^n$. 我们要做的是把我们对 $d\alpha/dt=T_t(\alpha)+g(t)$ 的解公式 $K_t\int_0^t K_s^{-1}(g(s))ds$ 转换到现在的背景下. 因为

$$\psi: f\ \mapsto\ \prec f,f',\cdots,f^{(n-1)}\succ$$

是从 n 阶方程 $L(f)=0$ 的解空间 $\boldsymbol{N}$ 到等价的一阶方程组 $d\boldsymbol{x}/dt=T_t(\boldsymbol{x})$ 的解空间 $\boldsymbol{N}$ 的一个同构, 从而得出, 如果我们有 N 的一组基 $\{u_i\}_1^n$, 则矩阵 $w_{ij}=u_j^{(i-1)}$ 的列构成了 $\boldsymbol{N}$ 的一组基 .

设 $\boldsymbol{w}(t)$ 为矩阵 $w_{ij}(t)=u_j^{(i-1)}(t)$. 由于在 t 的取值是从 $\boldsymbol{N}$ 到 $\mathbb{R}^n$ 的同构, 则 $\boldsymbol{w}(t)$ 的列对每个 t 构成了 $\mathbb{R}^n$ 的一组基. 但是 $K_t(\boldsymbol{\alpha})$ 是方程 $d\boldsymbol{x}/dt=T_t(\boldsymbol{x})$ 的通过初值点 $\prec 0,\boldsymbol{\alpha}\succ$ 的解在 t 的值, 从而推导出线性变换 K_t 将矩阵 $\boldsymbol{w}(0)$ 带到 $\boldsymbol{w}(t)$ 相应的列. K_t 的矩阵从而是 $\boldsymbol{w}(t)\cdot\boldsymbol{w}(0)^{-1}$, 那么我们的公式

$$\boldsymbol{f}(t)=K_t\int_0^t (K_s)^{-1}(\boldsymbol{g}(s))ds$$

的矩阵形式便是

$$\boldsymbol{f}(t)=\boldsymbol{w}(t)\cdot\boldsymbol{w}(0)^{-1}\cdot\int_0^t \boldsymbol{w}(0)\boldsymbol{w}(s)^{-1}\cdot\boldsymbol{g}(s)ds\ .$$

加之, 由于积分和常值线性变换 (这里是乘以一个常数矩阵) 的作用互换, 则中间的 $\boldsymbol{w}(0)$ 因子被消去, 我们便有了结果:

$$\boldsymbol{f}(t)=\boldsymbol{w}(t)\int_0^t \boldsymbol{w}(s)^{-1}\cdot\boldsymbol{g}(s)ds$$

是 $d\boldsymbol{x}/dt=T_t(\boldsymbol{x})+\boldsymbol{g}(t)$ 通过 $\prec 0,\boldsymbol{0}\succ$ 的解. 最后, 令 $\boldsymbol{k}(s)=\boldsymbol{w}(s)^{-1}\cdot\boldsymbol{g}(s)$, 故此解公式分裂成了一对:

$$\boldsymbol{f}(t)=\boldsymbol{w}(t)\int_0^t \boldsymbol{k}(s)ds \quad \text{和} \quad \boldsymbol{w}(s)\cdot\boldsymbol{k}(s)=\boldsymbol{g}(s)\ .$$

现在我们解非齐次 n 阶方程 $L(f)=g$, 意味着解一阶方程组, 其中 $\boldsymbol{g}=\prec 0,\cdots,0,g\succ$. 因此, 上面的第二个方程等价于

$$\sum_j w_{ij}(s)k_j(s)=0,\quad i<n\ ,$$

$$\sum_j w_{nj}(s)k_j(s)=g(s)\ .$$

另外, n 阶方程的解 f 是 n 联组 $\boldsymbol{f}$ 的第一个分量 (即 $f=\psi^{-1}\boldsymbol{f}$), 故我们最终有

$$f(t)=\sum_{j=1}^n w_{1j}(t)\int_0^t k_j(s)ds=\sum_1^n u_j(t)c_j(t)\ ,$$

其中 $c_j(t)$ 是反导数 $\int_0^t k_j(s)ds$. 任何其他的反导数也同样可以, 这是因为这两个所得的最后公式之间的差具有形式 $\sum_1^n a_iu_i(t)$, 这是齐次方程 $L(f)=0$ 的一个解. 现在我们已经证明了下面的定理:

定理 5.1 如果 $\{u_i(t)\}_1^n$ 是齐次方程 $L(h)=0$ 的解空间的一个基, 并且如果 $f(t)=\sum_1^n c_i(t)u_i(t)$, 其中导数 $c_i'(t)$ 是作为方程

$$\sum_i c_i'(t)u_i^{(j)}(t)=0,\quad j=0,\cdots,n-2\ ,$$

$$\sum c_i'(t)u_i^{(n-1)}(t)=g(t)$$

的解被确定, 则 $L(f)=g$.

我们现在考虑这个方法的简单例子. 方程 $y''+y=\sec x$ 具常系数, 我们可容易地找出齐次方程 $y''+y$ 的零空间. 它的一个基是 $\{\sin x,\cos x\}$. 我们在这里不

能用待定系数法, 因为 sec$x$ 不是一个常系数方程的解. 因此我们试用一个解

$$v(x) = c_1(x)\sin x + c_2(x)\cos x\ .$$

我们要解的方程组是

$$c_1'\sin x + c_2'\cos x = 0$$

$$c_1'\cos x - c_2'\sin x = \sec x\ .$$

因此 $c_2' = -c_1'\tan x$ 和 $c_1'(\cos x + \sin x \cdot \tan x) = \sec x$ 给出

$$c_1' = 1, \quad c_2' = -\tan x\ ,$$

$$c_1 = x, \quad c_2 = \log\cos x\ ,$$

从而

$$v(x) = x\sin x + (\log\cos x)\cos x\ .$$

(验证它 !)

这就是我们关于求解过程要说的全部内容. 在一切都顺利的情形下我们便完全掌握了 $L(f) = g$ 的解, 从而可以解初值问题. 如果 L 的阶为 n, 那么我们知道零空间 N 便是 n 维的, 而如果对给出的 g, 函数 v 是非齐次方程 $L(f) = g$ 的一个解, 则所有解的集合是 n 维平面 (仿射子空间)$M = N + v$. 如果我们已经找到了 N 的一个基 $\{u_i\}_1^n$, 则 $L(f) = g$ 的每个解具有形式 $f = \sum_1^n c_iu_i + v$. 初值问题是求 f 使得 $L(f) = g$ 且 $f(t_0) = a_1^0$, $f'(t_0) = a_2^0, \cdots, f_{(t_0)}^{(n-1)} = a_n^0$, 其中 $\prec a_1^0, \cdots, a_n^0 \succ = \boldsymbol{a}^0$ 是给定的初值. 现在我们用这 n 个条件来确定在 $f = \sum c_iu_i + v$ 中的 n 个系数 c_i 从而找到这个惟一的 f. 我们得到 n 个未知量 c_i 的 n 个方程. 我们解此问题的能力再次惟一地回溯到矩阵 $w_{ij}(t_0) = u_j^{(i-1)}(t_0)$ 是非异的这个事实, 这同我们在完成参数变值法过程中取得成功的原因一样 .

我们以讨论一个非常简单但重要的例题来结束本节. 当一个具有完美弹性的弹簧被拉伸或压缩时, 它以一个与形变成比例的 "恢复" 力进行阻抗. 如果我们画出此线圈形弹簧处在沿 x 轴的位置, 其在未扰动时一端固定而自由端位于原点 (图 6.3), 于是当此线圈伸长了距离 x(压缩为负伸长), 它所释放的力为 $-cx$, 其中 c 为代表弹簧的强度或弹性的一个常数, 负号表示力的方向指向位移的相反方向. 这是*胡克定律* .

假设我们在弹簧自由端加上了一个质量为 m 的点, 并把弹簧拉到初始位置 $x_0 = a$, 然后松开. 读者完全熟知此系统将振动, 而我们想要显式地描述这个振

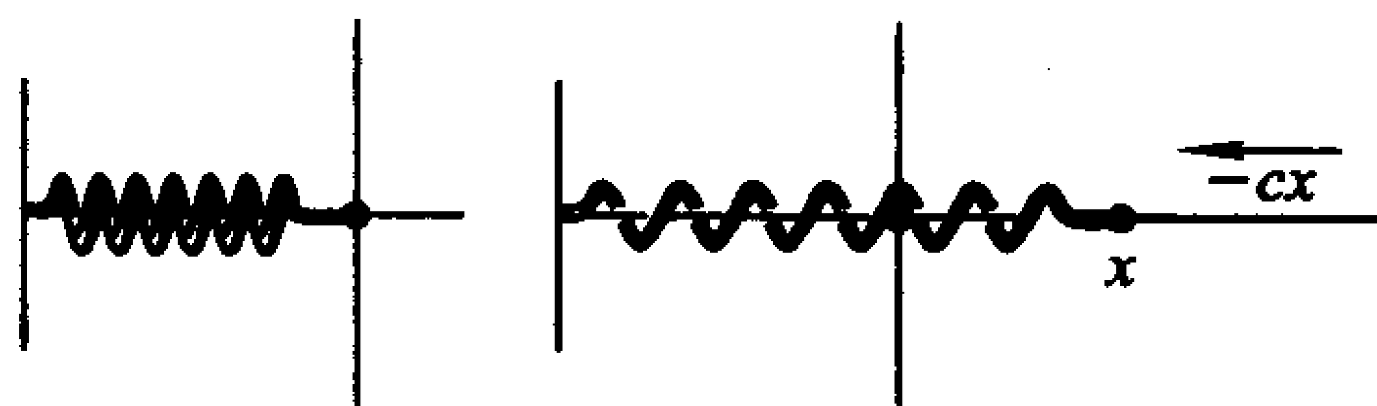


图 6.3

动. 我们忽略弹簧本身的质量 (等于修正 m) 而且暂且认为摩擦为零. 故此系统将不停地振动下去. 牛顿定律告诉我们, 如果力 F 作用于质量 m, 则此质点依照方程

$$m\frac{d^2x}{dt^2}=F\ .$$

加速. 这里的 $F=-cx$, 故此结合了牛顿定律和胡克定律的方程为

$$m\frac{d^2x}{dt^2}+cx=0\ .$$

这差不多是最简单的常系数方程了, 而且我们知道其通解为

$$z=c_1\sin\ \Omega t+c_2\cos\ \Omega t\ ,$$

其中 $\Omega=\sqrt{c/m}$. 我们的初始条件是当 $t=0$ 时 $x=a$ 和 $x'=0$. 因此 $c_2=a$, $c_1=0$, 故 $x=a\cos\ \Omega t$. 质点将永远在 $x=-a$ 与 $x=a$ 之间振动下去. 最大位移 a 被称作此振动的振幅 A. 单位时间的完全振动的次数被称作频率f, 故 $f=\Omega/2\pi=\sqrt{c}/2\pi\sqrt{m}$. 这是一个直观上清楚的事实的量化表达, 即频率随强度 c 增加而增加但随质量 m 增加而减小. 其他的初始条件也同样合理. 我们可以考虑此系统原本处于静止状态而给予击打, 故开始时初速为 v 而初始位移为 0, 即 $t=0$ 的状态. 现在 $c_2=0$, 而 $x=c_1\sin\ \Omega t$. 为了给 c_1 定值, 我们记起 $t=0$ 时 $dx/dt=v$, 因 $dx/dt=c_1\Omega\cos\ \Omega t$, 我们有 $v=c_1\Omega$, 从而 $c_1=v/\Omega$. 这是此运动的振幅. 一般说来, 初始条件会是 $t=0$ 时 $x=a$ 和 $x'=v$, 而如此决定的惟一解包含了通解中的全部两个项, 其振幅可以被算出来 .

当摩擦力被考虑进来时, 这种情况更为真实也更为有趣. 摩擦阻力理想的是一个与速度 dx/dt 成比例的力但又具有一个负号, 这是因为它的方向与运动方向相反. 因此我们的新方程是

$$m\frac{d^2x}{dt^2}+k\frac{dx}{dt}+cx=0\ ,$$

而且我们知道这个系统将按照常数 m, k, c 之间的关系以十分不同的方式行动. 要求读者在习题中进一步探索这些方程 .

令人惊奇的是相同的这个方程恰好也支配了一个自由振荡电路. 该方程现在被写成

$$L\frac{d^2x}{dt^2}+R\frac{dx}{dt}+\frac{1}{C}x=0\ ,$$

其中 L, R, C 分别是此电路的电感, 电阻和电容, 而 dx/dt 是电流. 然而这样一个电路的通常运行所涉及的不是自由振荡而是强迫振荡. 一个交流 (正弦曲线的) 电压被作为一个额外的, 外部的 "力" 施加于此电路, 方程现在为

$$L\frac{d^2x}{dt^2}+R\frac{dx}{dt}+\frac{1}{C}x=a\sin\omega t\ .$$

这展示了最有兴趣的现象. 利用待定系数法, 我们发现这个解包含了由齐次方程贡献的瞬时项它逐渐消逝, 还包含了由非齐次项 $a\sin\omega t$ 产生的具有频率 $\omega/2\pi$ 的项, 它成了恒定的部分. 称作位相和共振的新现象出现了, 读者将在习题中发现这些 .

习题

求下列方程的特解

5.1 $x''-x=t^4$.

5.2 $x''-x=\sin t$.

5.3 $x''-x=\sin t+t^4$.

5.4 $x''+x=\sin t$.

5.5 $y''-y'=x^2$(此处 $y'=dy/dx$) .

5.6 $y''-y'=e^x$.

5.7 考虑在正文中解过的方程 $y''+y=\sec x$. 在什么区间 I 上我们必须限制我们的讨论? 验证正文中找到的特解是正确的. 解下面的初值问题：

$$f''(x)+f(x)=\sec x,\quad f(0)=1,\quad f'(0)=-1\ .$$

用参数变值法解下列方程 .

5.8 $x''+x=\tan t$.

5.9 $x'''+x'=t$.

5.10 $y'' + y = 1$.

5.11 $y^{(4)} - y = \cos x$.

5.12 $y'' + 4y = \sec 2x$.

5.13 $y'' + 4y = \sec x$.

5.14 证明无摩擦弹性方程 $m(d^2x/dt^2) + cx = 0$ 的通解

$$C_1\sin \Omega t + C_2\cos \Omega t$$

可以重写为形式

$$A\sin (\Omega t - \alpha) .$$

(记住 $\sin(x - y) = \sin x\cos y - \cos x\sin y$.) 沿着一条直线的这种类型的运动称作简谐运动 .

5.15 在上面的习题中用初值 $t = 0$ 时 $dx/dt = v$, $x = a$ 来表达 A 和 α .

5.16 现考虑自由振动系统, 但考虑了摩擦因素. 因此有方程

$$m(d^2x/dt^2) + k(dx/dt) + cx = 0 ,$$

其中所有系数为正. 证明如果 $k^2 < 4mc$, 则此系统将永远振动, 但振幅按指数递降. 决定振动频率. 用习题 5.14 化简此解并描绘其图像 .

5.17 证明如果摩擦力足够大 ($k^2 \geqslant 4mc$), 则一个自由振动系统实际上不会摆动. 取最简单情形 $k^2 = 4mc$, 概述当 $t = 0$ 时 $dx/dt = 0, x = a$ 的初始条件下此系统的行为. 对初始条件为 $t = 0$ 时 $dx/dt = v, x = 0$ 做同样的事 .

5.18 用待定系数法求驱动电路方程

$$L\frac{dx^2}{dt^2} + R\frac{dx}{dt} + \frac{x}{c} = a \sin \omega t$$

的一个特解. 假定 $R > 0$, 用一种一般论证方法证明你的特解实际上是通解的稳恒态部分 (即不是指数式衰减的部分) .

5.19 证明上面习题中的你的解的 "电流" dx/dt 可以写成形式

$$\frac{dx}{dt} = \frac{a}{\sqrt{R^2 + X^2}} \sin(\omega t - \alpha) ,$$

其中 $X = L\omega - 1/\omega C$. 称 α 为位相角 .

5.20 继续讨论. 证明当 "外加电压" $a\sin \omega t$ 的频率为 $1/2\pi\sqrt{LC}$ 时, 电路中的电流将有一个极大振幅. 这是共振现象. 又证明当电流与外加电压同位相 (即 $\alpha = 0$) 当且仅当 $L = C = 0$.

5.21 位相 α 近似于 90° 和 -90° 的条件是什么？

5.22 在动态系统的稳定平衡点的理论中我们最终有了两个纯量积 (ξ,η) 和 $((\xi,\eta))$, 它们定义在一个有限维向量空间 V 上, 二次形式 $q(\xi)=\frac{1}{2}((\xi,\xi))$ 为势能, 而 $p(\xi')=\frac{1}{2}(\xi',\xi')$ 为动能. 我们现在知道有 $dq_\alpha(\xi)=((\alpha,\xi))$, 相似地对 p 也成立, 并且由于这个事实可以证明其拉格朗日方程能写成

$$\frac{d}{dt}\left(\frac{d\xi}{dt},\eta\right)=((\xi,\eta))\ .$$

证明可以找到 V 的一个基 $\{\beta_i\}_1^n$ 使得此向量方程成为二阶方程组

$$\frac{d^2x_i}{dt^2}=\lambda_i x_i,\quad i=1,\cdots,n\ ,$$

其中常数 λ_i 为正. 因而表明此系统的运动是 n 个线性无关的简谐运动的和．

6.6　边值问题

我们现在转向一个看起来像是初值问题的问题, 但它原来却具有完全不同的特性. 假设 T 是个二阶算子, 我们在闭区间 $[a,b]$ 上考虑它. 在物理中一些最重要的问题要求我们求 $T(f)=g$ 的解, 使得 f 在 a 和 b 具有已知值而不是 f 和 f' 在单个点 t_0 具有已知值. 称这个新问题为*边值问题*.这是由于 $\{a,b\}$ 是区域 $[a,b]=I$ 的边界的缘故. 像初值问题那样, 如果集合

$$M=\{f\in\mathfrak{C}^2([a,b]):f(a)=f(b)=0\}$$

最后被证明是 T 的零空间 N 的补空间, 则边值问题也可简单地被分裂成两个子问题, 然而, 如果读者考虑一下这个一般性的问题, 他会意识到从我们对初值问题的阐述中他没有得到什么线索, 事实上不得不设计出一个全新的工具．

我们的做法是忘记我们在试图解边值问题, 取而代之的是从内积和自伴算子的观点来思索线性微分算子 T．那就是说, 我们目前对 T 的研究就是利用内积 $(f,g)=\int_a^b f(t)g(t)dt$ 求 T 的右逆 S, 以解 $Tf=g$ 的这个常常遇到的一般性问题. 也像平常那样, S 可以由求 $N(T)$ 的补 M 来决定. 然而现在得到的结果是, 如果 T 是 “形式地自伴”, 则 M 的适当选取将使其相伴的右逆 S 为自伴和紧的, 并且可以作为齐次方程 $Tf-rf=0$ 的那些在 M 中的解来计算 S 的特征向量, 从而利用第五章的定理 5.1, 它允许我们以同样容易 (相对地) 的方法处理 S, 它们让我们处在早先的有限维情形．

我们先考虑一个 n 阶线性微分算子 T 的"形式自伴"概念. 分部积分的通常公式是

$$\int_a^b f'g = fg\Big]_a^b - \int_a^b fg' .$$

它允许出现在内积 (Tf,g) 中的 f 的导数每次转移一个到 g 上. 最终, f 没有被微分而 g 却被某个 n 阶线性微分算子 R 作用. 像前面的 $fg]_a^b$ 那样, 在端点的取值经过一步一步的积累, 可以描述为

$$B(f,g)]_a^b = \sum_{0\leqslant i+j<n} k_{ij}(x)f^{(i)}(x)g^{(j)}(x)]_a^b ,$$

其中系数函数 $k_{ij}(x)$ 是系数函数 $a_i(x)$ 和它们的导数的线性组合. 因此

$$(Tf,g) = (f,Rg) + B(f,g)]_a^b .$$

称算子 R 为 T 的*形式伴随*, 并且如果 $R=T$, 我们则说 T 是*形式自伴*.

分部积分公式每用一次便引起了一个符号变化, 而读者应该能看出 R 的首项系数是 $(-1)^n$ 乘以 T 的首项系数. 假定如此, 我们便看出了形式自伴性的一个必要条件是 n 为偶数, 从而 R 和 T 有相同的首项.

假设 T 为形式自伴, 我们寻找 T 的零空间 N 在 $\mathfrak{C}^n([a,b])$ 中的一个补 M, 它具有进一步的性质, 即与 T 相关的右逆 S 是作为从准希尔伯特空间 $\mathfrak{C}^\circ([a,b])$ 到它自身的自伴映射. 让我们看一看这进一步的要求等价于什么. 对于任意 $u,v\in\mathfrak{C}^\circ$, 令 $f=Su, g=Sv$, 从而 f 和 g 在 M 中, 以及 $u=Tf, v=Tg$. 于是 $(u,Sv)=(Tf,g)=(f,Tg)+B(f,g)]_a^b=(su,v)+B(f,g)]_a^b$.

我们因而有:

引理 6.1 *如果 T 是一个形式自伴的微分算子, M 为 T 的零空间的一个补空间, 则由 M 决定的 T 的右逆是自伴的当且仅当*

$$f,g\in M \;\Rightarrow\; B(f,g)]_a^b = 0 .$$

从现在开始我们将只考虑二阶的情形. 然而我们将要做的所有一切对一般情形都完全可行, 一般化的代价只不过增加了符号的复杂性.

我们从计算二阶算子 $Tf = c_2f'' + c_1f' + c_0f$ 的形式伴随开始, 我们有

$$(Tf,g) = \int_a^b c_2f''g + \int_a^b c_1f'g + \int_a^b c_0fg \ ,$$

$$\int_a^b c_1f'g = c_1fg]_a^b - \int_a^b f(c_1g)' \ ,$$

$$\begin{aligned}\int_a^b c_2f''g &= c_2f'g]_a^b - \int_a^b f'(c_2g)' \\ &= (c_2f'g - f(c_2g)')]_a^b + \int_a^b f(c_2g)'' \ ,\end{aligned}$$

给出了

$$(f,Rg) = \int f(c_2g)'' - (c_1g)' + (c_0g) \ ,$$

和

$$B(f,g) = c_2(f'g - g'f) + (c_1 - c_2')fg \ .$$

因此 $Rg = c_2g'' + (2c_2' - c_1)g' + (c_2'' - c_1' + c_0)g$, 而 $R = T$ 当且仅当 $2c_2' - c_1 = c_1$(和 $c_2'' - c_1' = 0$), 因此 $c_2' = c_1$. 我们已经证明了

引理 6.2 *二阶微分算子 T 为形式自伴当且仅当*

$$Tf = c_2f'' + c_2'f' + c_0f = (c_2f')' + c_0f \ ,$$

此时

$$B(f,g) = c_2(f'g - g'f) \ .$$

因此, 一个常系数算子为形式自伴当且仅当 $c_1 = 0$.

假设 T 为形式自伴, 我们现在想找出它的零空间的补 M 使得 $f, g \in M \Rightarrow B(f,g)]_a^b = 0$. 因为 N 是二维的, 故任意的补空间 M 可以由在 $X_2 = \mathfrak{C}^2([a,b])$ 中两个线性泛函 l_1 和 l_2 的零空间的交来描述. 例如在前面与初值问题有关的那个"单点" 补空间 M_{t_0} 便是两个线性泛函 $l_1(f) = f(t_0)$ 和 $l_2(f) = f'(t_0)$ 零空间的交. 但是在这里的两个 l_1 和 l_2 对两个函数 f 和 g 为零必须要使得 $B(f,g)]_a^b = c_2(f'g - g'f)]_a^b = 0$, 因而 $l_i(f)$ 必定包含了 f 和 f' 在 a 和 b 的值. 我们自然会猜测, l_1 和 l_2 中每个必定具有形式 $l(f) = k_1f(a) + k_2f'(a) + k_3f(b) + k_4f'(b)$, 事实上可以证明它的确如此 .

我们的问题现在可以重述如下. 我们必须找出两个线性泛函 l_1 和 l_2, 它们具有上面的那种一般形式, 使得如果 M 为它们的零空间的交, 则

(a) M 为 N 的一个补空间, 而且

(b) $f,g \in M \Rightarrow c_2(f'g-g'f)]_a^b = 0$,

这时, 我们称边界条件 $l_1(f) = l_2(f) = 0$ 是自伴的 .

引理 6.3 上面的条件 (a) 可以换为

(a′) T 在 M 上为单射的 .

证明 如果 T 在 M 上为单射的, 则 $M \cap N = \{0\}$, 故映射

$$f \mapsto \prec l_1(f), l_2(f) \succ$$

在 N 上为单射的, 并且由于 N 是二维的, 它应是从 N 到 $\mathbb{R}^2$ 的一个同构 (第二章定理 2.4 的推论). 于是由第一章的定理 5.3 知 M 是 N 的一个补空间 . □

现在我们可以容易地写出许多对的 l_1 和 l_2, 它们构成了一个自伴的边界条件. 我们在下面列出其中几个 .

(1) $f \in M \Leftrightarrow f(a) = f(b) = 0$[即 $l_1(f) = f(a), l_2(f) = f(b)$] .

(2) $f \in M \Leftrightarrow f'(a) = f'(b) = 0$.

(3) 更一般地, $f'(a) = kf(a), f'(b) = cf(b)$. (事实上, l_1 可以是任意一个只依赖于在 a 的值的泛函 l, 而 l_2 只依赖于 b. 因此 $l_1(f) = k_1 f(a) + k_2 f'(a)$, 从而如果 $l_1(f) = l_1(g) = 0$, 则对 $\prec f(a), f'(a) \succ$, $\prec g(a), g'(a) \succ$ 是相关的, 都在 l_1 的一维零空间中, 故 $f'g - g'f|_a = 0$. 对 l_2 和 b 成立相同的结果, 故此分离的一对端点条件使得 B 在 a 和 b 的值分别为 0, 从而 $B(f,g)|_a^b = 0$.)

(4) 如果 $c_2(a) = c_2(b)$, 则取 $f \in M \Leftrightarrow f(a) = f(b)$ 及 $f'(a) = f'(b)$, 就是说 $l_1(f) = f(a) - f(b)$ 和 $l_2(f) = f'(a) - f'(b)$.

我们要证明在除去 (3) 以外的每种情形中, 如果我们用 $T - \lambda$ 替换 T, 则条件 (a′) 仍然成立; 这里的 λ 为适当的数. 对情形 (3) 这个事实也是对的, 只是证明有些难, 故略去 .

引理 6.4 假设 M 由上面 (1),(2) 或 (4) 中任一自伴边界条件所定义, 并假设在 $[a,b]$ 上 $c_2(t) \geqslant m > 0$ 和 $\lambda \geqslant c_0(t) + 1$. 于是, 对所有 $f \in M$ 有

$$|((T-\lambda)f, f)| \geqslant m\|f'\|_2^2 + \|f\|_2^2 \,.$$

特别地, M 是 $T-\lambda$ 的零空间的一个补空间, 因而定义了 $T-\lambda$ 的一个自伴右逆.

证明 我们有

$$
\begin{aligned}
((\lambda-T)f,f) &= -\int_a^b (c_2f')'f+\int_a^b(\lambda-c_0)f^2 \\
&= -c_2f'f]_a^b+\int_a^b c_2(f')^2 \\
&\quad +\int_a^b(\lambda-c_0)f^2 .
\end{aligned}
$$

在条件 (1),(2) 或 (4) 中任一个下, $c_2f'f]_a^b=0$, 而且这两个积分项显然分别以 $m\|f'\|_2^2$ 和 $\|f\|_2^2$ 为下界. 于是引理 6.3 表明 M 是 $T-\lambda$ 的零空间的补 . □

现在要讲我们的主要的定理了. 它说的是, 由上面子空间 M 决定的 $T-\lambda$ 的右逆 S 是准希尔伯特空间 $\mathfrak{C}^\circ([a,b])$ 到自身的一个紧自伴映射, 因此被赋予了上一章定理 5.1 的全部充足的特征值结构. 首先我们提及 n 个经典术语. $[a,b]$ 上的一个*施图姆 - 刘维尔系统*是定义在闭区间 $[a,b]$ 上的一个形式自伴二阶微分算子 $Tf=(c_2f')'+c_0f$, 连同对此区间的一个自伴边界条件 $l_1(f)=l_2(f)=0$. 如果 $c_2(t)$ 在 $[a,b]$ 上永不为零, 则称此系统是*正则的*. 如果 $c_2(a)$ 或 $c_2(b)$ 为零, 或者如果区间 $[a,b]$ 被换作诸如 $[a,\infty)$ 这样的无穷区间, 则称此系统为*奇异的* .

定理 6.1 *如果 $T:l_1,l_2$ 是 $[a,b]$ 上一个正则施图姆 - 刘维尔系统, 其中 $c_2>0$, 则如果取 λ 充分大, 由齐次边界条件定义的子空间 M 是 $N(T-\lambda)$ 的补空间, 从而由 M 决定的 $T-\lambda$ 的右逆是从准希尔伯特空间 $\mathfrak{C}^\circ([a,b])$ 到自身的一个紧的自伴映射 .*

证明 此证明有赖于前面引理中的不等式. 由于我们仅在边界条件 (1),(2) 和 (4) 之下证明了这个不等式, 那么对此定理的证明也只对这些情形有效 .

令 $g=(T-\lambda)f$. 因为由施瓦兹不等式有 $\|g\|_2\|f\|_2\geqslant|((T-\lambda)f,f)|$, 那么由引理我们首先有 $\|f\|_2^2\leqslant\|g\|_2\|f\|_2$, 使得

$$\|f\|_2\leqslant\|g\|_2 ,$$

其次有 $m\|f'\|_2^2\leqslant\|g\|_2\|f\|_2\leqslant\|g\|_2^2$, 故

$$\|f'\|_2\leqslant\|g\|_2/\sqrt{m} .$$

我们已经验证过了由 M 定义的形式自伴 $T-\lambda$ 的右逆 S 是自伴的, 我们还需证明的是集合 $S[U]=\{f:\|g\|_2\leqslant1\}$ 具有紧的闭包. 对任何一个这样的 f, 施

瓦兹不等式和上面的不等式表明

$$|f(y)-f(x)| \leqslant \int_x^y |f'| = \int_x^y |f'|\cdot 1 \leqslant \|f'\|_2\ |y-x|^{1/2}$$
$$\leqslant \frac{|y-x|^{1/2}}{\sqrt{m}}.$$

因此 $S[U]$ 为一致等度连续. 因为 $S[U]$ 中函数的公共定义域为紧集 $[a,b]$, 那么, 如果我们能证明存在一个常数 C, 使得 $S[U]$ 中所有函数的值域都在 $[-C,C]$ 中, 我们则可从第四章的定理 6.1 得出结论说集合 $S[U]$ 是完全有界的. 在上面的最后一个不等式中取 $|f|$ 达到其极大值和极小值的点 y 和 x, 我们便有了 $\|f\|_\infty - \min|f| \leqslant (b-a)^{1/2}/\sqrt{m}$. 但是 $(\min|f|)(b-a)^{1/2} \leqslant \|f\|_2 \leqslant \|g\|_2 \leqslant 1$, 因此

$$\|f\|_\infty \leqslant C = 1/(b-a)^{1/2} + (b-a)^{1/2}/\sqrt{m}.$$

因此 $S[U]$ 是由紧集 $[a,b]$ 到紧集 $[-C,C]$ 的函数映射的一致等度连续集合, 从而在一致范数下为完全有界. 因为 $\mathfrak{C}([a,b])$ 在一致范数下是完备的, 所以 $S[U]$ 中每个序列都有一个子序列一致收敛于某个 $f \in \mathfrak{C}$, 又因为 $\|f\|_2 \leqslant (b-a)^{1/2}\|f\|_\infty$, 故此子序列在 2 - 范数下也收敛于 f.

那么, 我们已经证明了如果 H 是有标准的内积的准希尔伯特空间 $\mathfrak{C}([a,b])$, 则单位球 $U \subset H$ 在 S 下的像 $S[U]$ 具有的性质是: $S[U]$ 中每个序列都有一在 H 中收敛的子序列. 这是我们实际用来证明第五章定理 5.1 的那个性质. 但它不完全是 S 紧性的定义, 后者要求我们证明闭包 $\overline{S[U]}$ 在 H 中紧. 但是, 如果 $\{\xi_n\}$ 为在此闭包中的任一序列, 那么我们可以在 $S[U]$ 中选取 $\{\zeta_n\}$ 使得 $\|\xi_n-\zeta_n\| < 1/n$. 由上面知序列 $\{\zeta_n\}$ 有一收敛子序列 $\{\zeta_{n(m)}\}_m$, 于是 $\{\xi_{n(m)}\}_m$ 收敛于同一极限. 因此 S 是个紧算子. □

定理 6.2 *存在完全由 T 的特征向量形成的法正交序列 $\{\varphi_n\}$, 它们是 M 的一组基. 另外, 任意 $f \in M$ 关于基 $\{\varphi_n\}$ 的傅里叶展开式一致收敛于 f(在 2 - 范数下同样也对).*

证明 由第五章定理 5.1 知, 对 S 的值域存在一个在 M 中的特征基. 因为对某个非零的 r_n 有 $S\varphi_n = r_n\varphi_n$, 故我们有 $(T-\lambda)(r_n\varphi_n) = \varphi_n$ 和 $T\varphi_n = ((1+\lambda r_n)/r_n)\varphi_n$. 此级数收敛的一致性来自下面更一般的考量. □

引理 6.5 *假设 T 为准希尔伯特空间 V 上的一个自伴算子, 并且作为从 V 到 $\prec V,q \succ$ 的映射 T 是紧的, 这里的 q 是 V 上的第二个范数, 它支配内积范数*

p(即 $q \geqslant cp$). 于是 T 是紧的 (从 p 到 p), 并且在 T 的值域中任一个元素 β 的特征基展开式 $\sum b_n\varphi_n$ 在这两个范数下都收敛于 β.

证明 设 U 为 V 中在内积范数下的单位球. 由此引理的假设条件, $T[U]$ 的 q - 闭包 B 是紧的. 于是 B 也是 p - 紧的, 这是因为由于 $p \leqslant cq$, 则其中的任何一个序列有一个 q - 收敛的子序列也必定 p - 收敛于同一极限. 因此可以应用特征基定理.

现在设 α 及 $\beta = T(\alpha)$ 分别有傅里叶级数 $\sum a_i\varphi_i$ 和 $\sum b_i\varphi_i$, 并设 $T(\varphi_i) = r_i\varphi_i$. 于是 $b_i = r_ia_i$, 因为 $b_i = (T(\alpha), \varphi_i) = (\alpha, T(\varphi_i)) = (\alpha, r_i\varphi_i) = r_i(\alpha, \varphi_i) = r_ia_i$. 鉴于部分和 $\sum_1^n a_i\varphi^i$ 的序列是 p - 有界的 (贝塞尔不等式), 那么序列 $\{\sum_1^n b_i\varphi_i\} = \{T(\sum_1^n a_i\varphi_i)\}$ 为完全 q - 有界. 因此它的任一子序列均有一个子子序列 q - 收敛于 V 中某个元素 γ. 由于它也 p - 收敛于 γ, γ 必定就是 β. 因此每个子序列有一个子子序列 q - 收敛于 β, 故而 $\{\sum_1^n b_i\varphi_i\}$ 本身由第四章引理 4.1, q - 收敛于 β. □

习题

6.1 已知 $Tf(x) = xf''(x) + f(x)$ 和 $Sf(x) = f'(x)$, 计算 $T \circ S$ 和 $S \circ T$.

6.2 证明微分算子 $T = aD$ 和 $S = bD$ 可交换当且仅当函数 $a(x)$ 和 $b(x)$ 成比例.

6.3 证明微分算子 $T = aD^2$ 和 $S = bD$ 可交换当且仅当 $b(x)$ 是一次多项式 $b(x) = cx + d$, $a(x) = k(b(x))^2$.

6.4 计算下列 T 的形式伴随 S:

(a) $Tf = f'$, (b) $Tf = f''$, (c) $Tf = f'''$,

(d) $(Tf)(x) = xf'(x)$, (e) $(Tf)(x) = x^3f''(x)$.

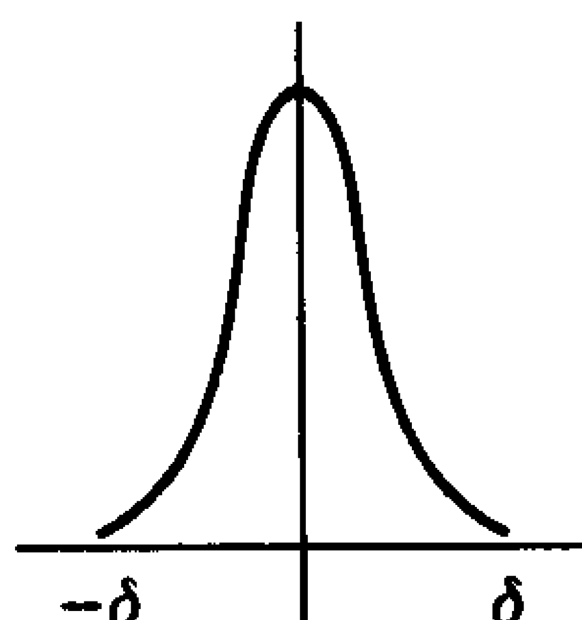
6.5 设 S 和 T 分别为 m 阶和 n 阶的线性微分算子. 要使 $S \circ T$ 为 $m + n$ 阶的线性微分算子, 它们的系数条件是什么?

6.6 设 T 是二阶线性微分算子

$$(Tf)(t) = a_2(t)f''(t) + a_1(t)f'(t) + a_0(t)f(t).$$

使其形式伴随存在的系数函数的条件是什么? 对 T 为 n 阶, 这些条件又是什么?

6.7 设 S 和 T 分别为 m 和 n 阶的微分算子, 并设所有系数为 $\mathfrak{C}^\infty$ 函数 (无限次可微). 证明 $S \circ T - T \circ S$ 的阶 $\leqslant m + n - 1$.



6.8 一个 δ - 反射脉冲是一个连续的非负函数 φ 使得在 $[-\delta,\delta]$ 外面为 0, 并且 $\int_{-\delta}^{\delta}\varphi=1$(图 6.4). 我们假定有一个无限可微的 1 - 反射脉冲 φ. 证明对每个 $\delta>0$ 存在一个无限可微的 δ - 反射脉冲. 定义你所想到的中心在 x 的一个 δ - 反射脉冲的意思并证明存在这样的一个 .

6.9 设 f 为 $[a,b]$ 上的连续函数使得 $(f,g)=\int_a^b fg=0$, 其中 g 为每个在 a 和 b 附近为 0 的无限可微函数. 证明 $f=0$. (利用上一个习题.)

6.10 设 $\mathfrak{C}^\infty([a,b])$ 为 $[a,b]$ 上无限次可微函数的向量空间, 并设 T 为系数在 $\mathfrak{C}^\infty$ 的二阶线性微分算子 :

$$(Tf)(t)=a_2(t)f''(t)+a_1(t)f'(t)+a_0(t)f(t)\ .$$

设 S 为 $\mathfrak{C}^\infty([a,b])$ 上的一个线性算子使得

$$(Tf,g)-(f,Sg)=K(f,g)$$

为只依赖于 f,g,f' 和 g' 在 a 和 b 的值的双线性泛函. 证明 S 是 T 的形式伴随.[提示: 取 f 为中心在 x 的 δ - 反射脉冲. 于是 $K(f,g)=0$. 现力图将此论断转成可以应用上面习题的形式以得到证明 .]

6.11 证明上面习题的 n 阶情形的推广形式 .

6.12 设 X 为具 $\mathfrak{C}^\infty$ 系数的线性微分算子的空间, 并设 A_T 为 T 的形式伴随. 证明 $T\mapsto A_T$ 是由 X 到 X 的同构. 证明 $A_{(T\circ S)}=A_S\circ A_T$.

6.7 傅里叶级数

没有多少正则施图姆 - 刘维尔系统的相伴法正交特征基被证明对实际的计算是重要的. 被使用的大多数的法正交基, 诸如属于贝塞尔的, 勒让德的, 埃尔米特的以及拉盖尔的, 全都由奇异施图姆 - 刘维尔系统产生, 因而超过了我们为此讨论所设定的界限. 然而最著名的例子, 傅里叶级数, 对我们来说倒是现成的 .

我们将考虑常系数算子 $Tf=D^2f$, 它或是在 $[0,\pi]$ 上的边界条件 $f(0)=f(\pi)=0$(类型 1) 或是在 $[-\pi,\pi]$ 上的周期边界条件 $f(-\pi)=f(\pi)$, $f'(-\pi)=f'(\pi)$(类型 4) 下既是形式自伴的又是正则的 .

要解第一个问题, 我们必须求 $f''-\lambda f=0$ 满足 $f(0)=f(\pi)=0$ 的解. 如果 $\lambda>0$, 我们知道其二维解空间由 $\{e^{rx},e^{-rx}\}$ 张成, $r=\lambda^{1/2}$. 但是如果 $c_1e^{rx}+c_2e^{-rx}$ 在 0 和 π 都为 0, 则 $c_1=c_2=0$ (因为 $\prec 1,1\succ$ 和 $\prec e^{r\pi},e^{-r\pi}\succ$ 线性

无关). 因此当 $\lambda > 0$ 时没有满足此边界条件的解. 如果 $\lambda = 0$, 则 $f(x) = c_1x + c_0$, 仍是 $c_0 = c_1 = 0$.

如果 $\lambda < 0$, 则这个解空间由 $\{\sin rx, \cos rx\}$ 张成, 其中 $r = (-\lambda)^{1/2}$. 如果 $c_1\sin rx + c_2\cos rx$ 在 $x = 0$ 和 $x = \pi$ 时为 0, 我们得到 $c_2 = 0$ 以及 $r\pi = n\pi$, n 为某个整数. 因此第一个系统的特征函数形成了集合 $\{\sin nx\}_1^\infty$, 而相应的 D^2 的特征值为 $\{-n^2\}_1^\infty$.

在本节最后部分我们将证明在 $\mathfrak{C}^2([a,b])$ 中那些在 a 和 b 附近为零的函数集在 2 - 范数下稠于 $\mathfrak{C}([a,b])$. 假定已知这个结果, 那么从第五章定理 2.3 可以推导出 M 的一组基是 $\mathfrak{C}^\circ$ 的一组基的结论, 从而我们有了下面的施图姆 - 刘维尔定理的推论 :

定理 7.1 序列 $\{\sin nx\}_1^\infty$ 是准希尔伯特空间 $\mathfrak{C}^\circ([0,\pi])$ 的一组正交基. 如果 $f \in \mathfrak{C}^2([0,\pi])$ 且 $f(0) = f(\pi) = 0$, 则 f 的傅里叶级数一致收敛于 f .

现在考虑第二个边界问题. 计算有一点复杂, 但仍然当 $f(x) = c_1e^{rx} + c_2e^{-rx}$, 且 $f(-\pi) = f(\pi), f'(-\pi) = f'(\pi)$ 时我们有 $f = 0$. 因为这时

$$c_1e^{-r\pi} + c_2e^{r\pi} = c_1e^{r\pi} + c_2e^{-r\pi} ,$$

给出 $c_1 = c_2$, 另外

$$c_1re^{-r\pi} - c_2re^{r\pi} = c_1re^{r\pi} - c_2re^{-r\pi} ,$$

给出了 $c_1(e^{r\pi} - e^{-r\pi}) = 0$, 故 $c_1 = 0$. $f(x) = c_1x + c_2$ 也须剔除. 最后, 如果 $f(x) = c_1\sin rx + c_2\cos rx$, 我们的边界条件变为

$$2c_1\sin r\pi = 0 \qquad \text{和} \quad 2rc_2\sin r\pi = 0 ,$$

故又有 $r = n$, 但这次, $(D^2 + n^2)f = 0$ 的整个解空间都满足这个边界条件 .

定理 7.2 集合 $\{\sin nx\}_1^\infty \cup \{\cos nx\}_0^\infty$ 形成了准希尔伯特空间 $\mathfrak{C}^\circ([-\pi,\pi])$ 的一个正交基. 如果 $f \in \mathfrak{C}^2([-\pi,\pi])$ 且 $f(-\pi) = f(\pi)$, $f'(-\pi) = f'(\pi)$, 则 f 的傅里叶级数在 $[-\pi,\pi]$ 上一致收敛于 f .

剩余的证明 除了 $\sin nx$ 和 $\cos nx$ 的正交性外, 此定理由一般的施图姆 - 刘维尔问题的讨论中得到. 我们有

$$\begin{aligned}(\sin nx, \cos nx) &= \int_{-\pi}^{\pi} \sin nt \cos nt\, dt \\ &= \frac{1}{2}\int_{-\pi}^{\pi} \sin 2nt\, dt \\ &= -(1/4n)\cos 2nx]_{-\pi}^{\pi} \\ &= 0\,.\end{aligned}$$

或者我们直接注意到第一个积分项是个奇函数, 因而它在任意对称区间 $[-a, a]$ 上的积分必为零.

具有不同特征值的特征向量的正交性自然来自第五章定理 3.1 的证明.

最后, 我们证明前面所需要的那个稠密性定理. 有非常巧妙的方法来证明它, 但这需要比我们现在掌握的更多的工具和技术; 与其花时间来打造这些机制还不如白手起家来证明此定理.

在表示一个函数类的记号放上一个零下标是一种标准符号, 它表示在这个类中取出的那些 "在边界上" 按某种意义为零的那些函数的函数类. 这里 $\mathfrak{C}_0([a,b])$ 就表示在 $\mathfrak{C}([a,b])$ 中那些在 a 和 b 的邻域中为零的函数, 类似地有 $\mathfrak{C}_0^2([a,b])$.

定理 7.3 *在一致范数下 $\mathfrak{C}^2([a,b])$ 稠于 $\mathfrak{C}([a,b])$ 而在 2 - 范数下 $\mathfrak{C}_0^2([a,b])$ 稠于 $\mathfrak{C}([a,b])$.*

证明 我们首先以一个逐段 "线性" 函数 g 以 ε 范围内逼近 $f \in \mathfrak{C}([a,b])$, 其中 g 是把 $[a,b]$ 的剖分 $a = x_0 < x_1 < \cdots < x_n = b$ 上的 f 图像上的相邻点用直线段相联结形成的函数. 如果在每个区间 (x_{i-1}, x_i) 上 f 的变化小于 ε, 则 $\|f - g\|_\infty \leqslant \varepsilon$. 现在 $g'(t)$ 是一个阶梯函数, 它在上面剖分的子区间上为常值. 我们在 $g'(t)$ 的每个跳跃之处稍稍改变一点使得到的新函数 $h(t)$ 在这些地方连续. 如果我们如图 6.5 所描绘的那样去做, 则在跳跃处的总的积分误差为零, $\int_{x_i-\delta}^{x_i+\delta}(h - g') = 0$, 而极大误差 $\int_{x_i-\delta}^{x_i}$ 是 $\delta\Delta/4$. 如果我们取 $\delta = \varepsilon/\|g'\|_\infty$, 则由于 $\Delta \leqslant 2\|g'\|_\infty$, 这个极大误差小于 ε. 我们现在有了一个连续函数 h, 满足 $|\int_a^x h(t)dt - (f(x) - f(a))| < 2\varepsilon$. 换句话说, 我们有一个一致逼近于 f 的连续可微函数.

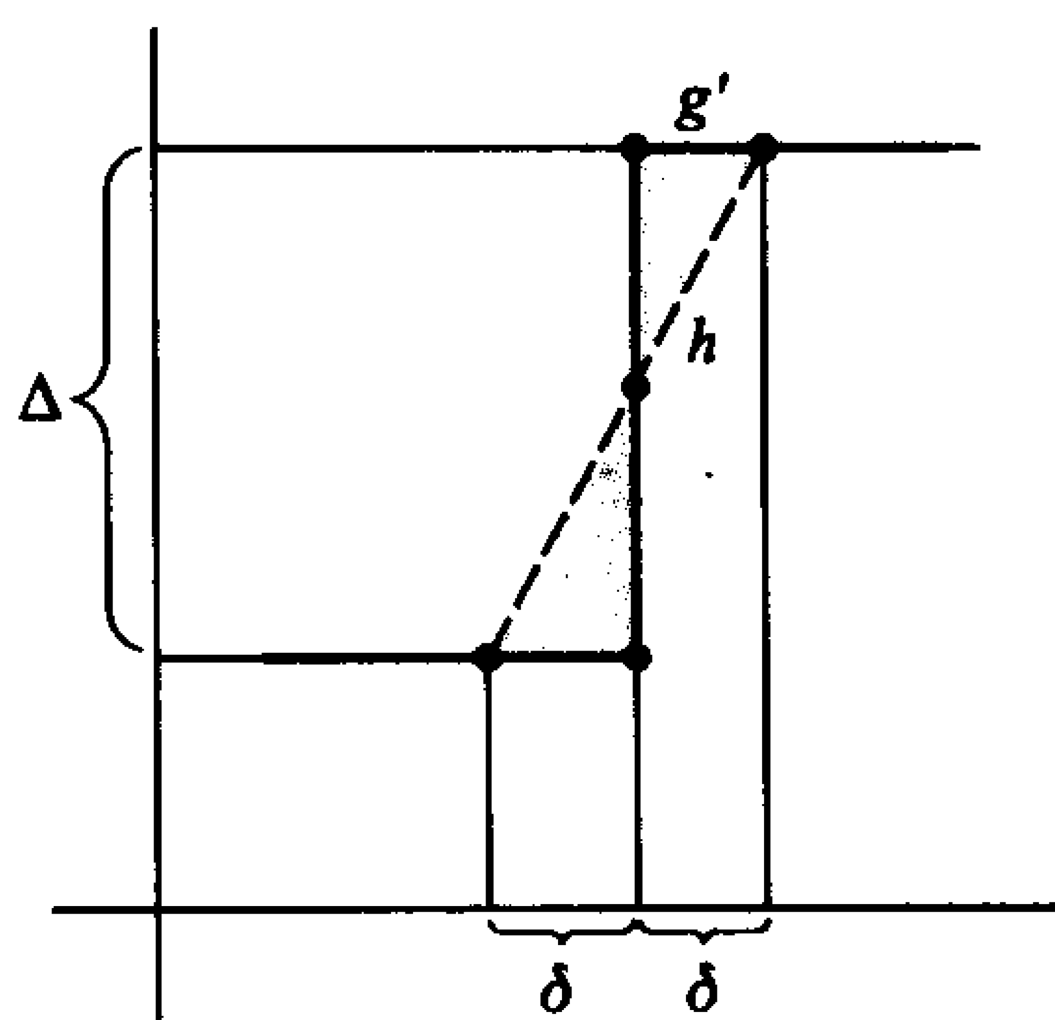


图 6.5

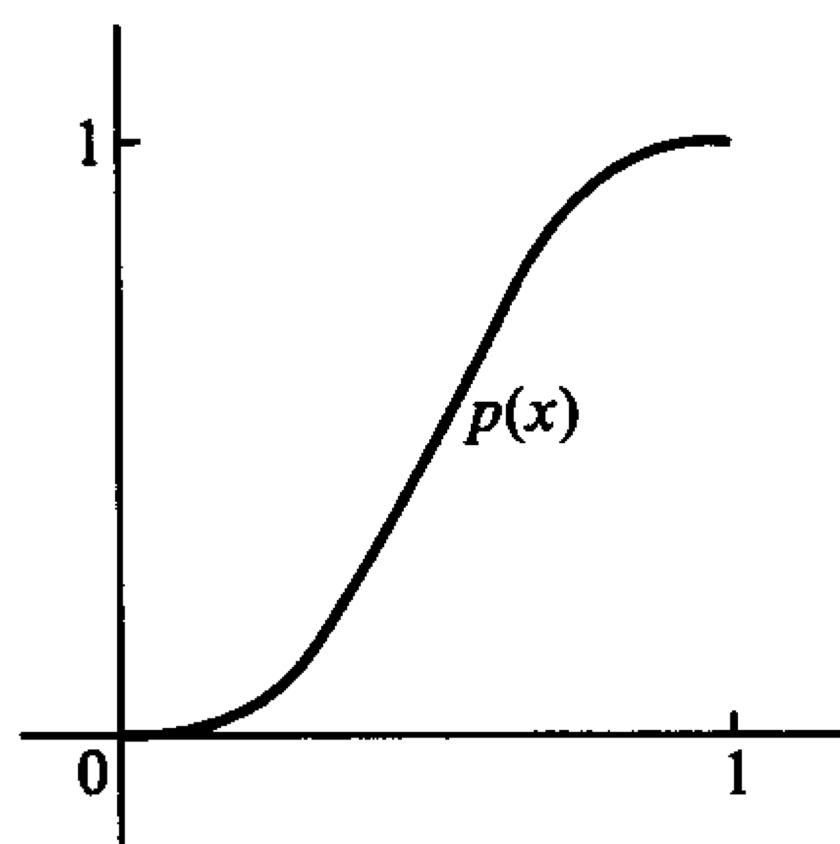


图 6.6

现在 $\mathfrak{C}^1([a,b])$ 中选取 g 和 h 使得 $\|f-g\|_\infty<\varepsilon/2$, 于是 $\|g'-h\|_\infty<\varepsilon/2(b-a)$. 那么

$$\left|g(x)-g(0)-\int_0^x h\right|<\varepsilon/2\ ,$$

从而 $H(x)=\int_0^x h+g(0)$ 是二次连续可微函数, 使得 $\|f-H\|_\infty<\varepsilon$. 换句话说 $\mathfrak{C}^2([a,b])$ 在 $\mathfrak{C}([a,b])$ 中按一致范数稠密. 因为

$$\|f\|_2=\left(\int_a^b f^2\right)^{1/2}\leqslant\|f\|_\infty\left(\int_1^b 1\right)^{1/2}=(b-a)^{1/2}\|f\|_\infty\ ,$$

所以它也在 2 - 范数下稠密 .

然而我们现在可以做对一致范数不可能做的一些事了: 我们能改变此近似函数为一个在 a 和 b 的邻域为 0 并保持好的 2 - 范数近似. 已知 δ, 设 $e(t)$ 为 $[a,b]$ 上一个非负函数使得 $e(t)=1$, 其中 $t\in[a+2\delta,b-2\delta]$ 而在 $[a,a+\delta]$ 和在 $[b-\delta,b]$ 上 $e(t)=0$, 同时 e'' 连续, $\|e\|_\infty=1$. 这样的 $e(t)$ 是存在的, 因为我们可以画出一个来. 我们把实际上定义一个 $e(t)$ 当作一个有意思的习题留给读者. 给一个提示: 以某种方式证明存在一个 5 次多项式 $p(t)$, 它在 0 和 1 之间的图像如图 6.6 所示, 并在 0 和 1 有一个为 0 的二阶导数, 然后利用此图像中的一段, 经适当的平移, 压缩, 旋转等等, 以便拼合成 $e(t)$.

无论如何, 对 $\mathfrak{C}([a,b])$ 上任意的 g 有 $\|g-eg\|_2\leqslant\|g\|_\infty(4\delta)^{1/2}$, 并且如果 g 有直到 2 阶的连续导数, 那么 eg 也如此. 因此, 如果我们从 $f\in\mathfrak{C}$ 开始, 以 $\mathfrak{C}^2$ 中的 g 近似它, 然后用 eg 近似 g, 把它们合起来便有了定理中的第二个逼近 . □

习题

7.1 把准希尔伯特空间 $\mathfrak{C}([0,\pi])$ 的正交基 $\{\sin nx\}_1^\infty$ 转化为一个法正交基 .

7.2 对 $\mathfrak{C}([-\pi,\pi])$ 上的正交基 $\{\sin nx\}_1^\infty \cup \{\cos nx\}_0^\infty$ 做上面同样的事 .

7.3 证明 $\{\sin nx\}_1^\infty$ 对于 $[-\pi,\pi]$ 上的所有奇连续函数是向量空间 V 一个正交基.(讨巧些. 不要从头算起.) 法化上面的基 .

7.4 叙述并证明对 $[-\pi,\pi]$ 上偶函数的相应定理 .

7.5 证明奇函数的导数为偶, 反之也对 .

7.6 我们现在要证明下面关于傅里叶级数一致收敛性的更强的定理 .

定理 7.4 *设 f 在 $[-\pi,\pi]$ 上有连续导数, 并假定 $f(-\pi)=f(\pi)$. 于是 f 的傅里叶级数一致收敛于 f .*

为方便起见, 假定 f 为偶函数 (这只是为了削减计算量.) 首先证明 f' 的傅里叶级数由对 f 的傅里叶级数经逐项微分得到. 把上面习题用于此处. 其次由它的傅里叶级数到 f' 的 2 - 范数收敛性和施瓦兹不等式证明 f 的傅里叶级数一致收敛 .

7.7 证明 $\{\cos nx\}_0^\infty$ 是 $[0,\pi]$ 上满足 $f'(0)=f'(\pi)=0$ 的 $\mathfrak{C}^2$ 函数的空间 M 的一个法正交基 .

7.8 求一个 5 次多项式 $p(x)$ 使得

$$p(0)=p'(0)=p''(0)=0,\quad p'(1)=p''(1)=0, p(1)=1 .$$

(在得到最后结果前先忘掉最后面的一个条件.) 描绘 p 的图像 .

7.9 用上面多项式 p 的 "一段" 构造函数 $e(x)$ 使得 e' 和 e'' 存在并连续, 在 $x<a+\delta$ 和 $x>b-\delta$ 时 $e(x)=0$, 而在 $[a+2\delta,b-2\delta]$ 上 $e(x)=1$, 且 $\|e\|_\infty=1$.

7.10 证明下面的在 $[0,\pi]$ 上的魏尔斯特拉斯定理, 其步骤如下. 我们知道 f 可以由一个 $\mathfrak{C}^2$ 函数 g 一致地逼近 .

(1) 证明可以找到 c 和 d, 使 $g(t)-c(t)-d$ 在 0 和 π 为 0 .

(2) 用此函数的傅里叶展开和对函数 $\sin nx$ 的马克劳林级数证明可以找到此多项式 $p(x)$.

定理 7.5 (魏尔斯特拉斯逼近定理) *在一致范数下多项式在 $\mathfrak{C}([a,b])$ 中稠密. 就是说, 给出 $[a,b]$ 上任意连续函数 f 和任意 ε, 存在一个多项式 p 使得对 $[a,b]$ 中所有 x 有 $|f(x)-p(x)|<\varepsilon$.*

第七章　多重线性泛函

本章主要是作参考用的. 虽然大部分证明都给出来了, 但并不期望读者去学它们. 我们的目标是一组关于交错多重线性泛函或外形式的基本定理, 行列式函数是我们得到的东西中的一个.

7.1　双线性泛函

我们已经学过双线性泛函的许多不同的方面. 在第一章第 6 节我们考虑了它们的对偶性所蕴涵的思想, 在第二章第 7 节考虑了对称双线性泛函的"典型形式"和它们的等价二次型, 当然, 第五章的整个内积理论是双线性泛函的一个更加特殊类的理论. 在本章中我们只局限于在有限维空间上的双线性和多重线性泛函, 我们所关心的是纯代数的.

我们从早先讲过的代数的有关材料开始. 如果 V 和 W 是有限维向量空间, 则在 $V \times W$ 上的所有双线性泛函显然是个向量空间. 我们以 $V^* \otimes W^*$ 表示它并称其为 V^* 和 W^* 的*张量积*. 我们的第一个定理只不过是叙述某些隐含在第一章定理 6.1 中的东西.

定理 1.1　*向量空间* $V^* \otimes W^*$, $\mathrm{Hom}(V, W^*)$ *和* $\mathrm{Hom}(W, V^*)$ *自然地同构*.

证明 在第一章定理 6 我们曾看到 $V^*\bigotimes W^*$ 中每个 f 决定了一个线性映射 $\alpha \mapsto f_\alpha : W$ 到 V^*, 其中 $f_\alpha(\xi) = f(\xi,\alpha)$, 我们还注意到从 $V^*\bigotimes W^*$ 到 $\mathrm{Hom}(W,V^*)$ 是双射的. 目前的这个定理全部所添加的只是说, 这个双射对应是线性的, 从而构成一个自然的同构, 对于 $V^*\bigotimes W^*$ 到 $\mathrm{Hom}(V,W^*)$ 的那个双射也是相似的. 要明白这点, 设 f_T 为对应于 $\mathrm{Hom}(V,W^*)$ 中 T 的双线性泛函. 于是由于 $f_{(T+S)}(\alpha,\beta) = ((T+S)(\alpha))(\beta) = (T(\alpha)+S(\alpha))(\beta) = (T(\alpha))(\beta)+(S(\alpha))(\beta) = f_T(\alpha,\beta)+f_S(\alpha,\beta)$, 则 $f_{(T+S)} = f_T + f_S$. 对于齐性可同样证明 .

$V^*\bigotimes W^*$ 与 $\mathrm{Hom}(W,V^*)$ 只要反转变量的角色就完全按同样的方法推出. 因此我们完成了证明 . □

在寻找 $V^*\bigotimes W^*$ 的基之前, 我们对任意两个泛函 $\gamma \in V^*$ 和 $\lambda \in W^*$ 定义一个双线性泛函 $\gamma\bigotimes\lambda$ 为 $(\gamma\bigotimes\lambda)(\xi,\eta) = \gamma(\xi)\lambda(\eta)$. 我们称 $\gamma\bigotimes\lambda$ 为泛函 γ 和 λ 的张量积, 并称任何具有此形式的双线性泛函为*初等的*. 不太难就可看出 $f \in V^*\bigotimes W^*$ 为初等当且仅当对应的 $T \in \mathrm{Hom}(V,W^*)$ 是个并向量 .

如果 V 和 W 是各为维数 m 和 n 的有限维向量空间, 则上述 $V^*\bigotimes W^*$ 与 $\mathrm{Hom}(V,W^*)$ 的同构表明 $V^*\bigotimes W^*$ 的维数是 mn. 我们现在用 V 和 W 中的给定基来描述它的基 .

定理 1.2 *设 $\{\alpha_i\}_1^m$ 和 $\{\beta_j\}_1^n$ 为 V 和 W 的任意基, 其在 V^* 和 W^* 中的对偶基为 $\{\mu_i\}_1^m$ 和 $\{\nu_j\}_1^n$. 则 mn 个初等双线性泛函 $\{\mu_i\bigotimes\nu_j\}$ 形成 $V^*\bigotimes W^*$ 的一组对应的基 .*

证明 因为 $\mu_i\bigotimes\nu_j(\xi,\eta) = \mu_i(\xi)\cdot\nu_j(\eta) = x_iy_j$, $f(\xi,\eta) = \sum_{i,j} t_{ij}x_iy_j$ 这个矩阵展式为 $f(\xi,\eta) = \sum_{i,j} t_{ij}(\mu_i\bigotimes\nu_j)(\xi,\eta)$ 或者写为

$$f = \sum_{i,j} t_{ij}(\mu_i \bigotimes \nu_j) .$$

于是集合 $\{\mu_i\bigotimes\nu_j\}$ 张成 $V^*\bigotimes W^*$. 因为它包含了与 $V^*\bigotimes W^*$ 维数 (mn) 相同个数的元素, 故是一组基 . □

当然, 也可直接验证其无关性. 如果 $\sum_{i,j} t_{ij}(\mu_i\bigotimes\nu_j) = 0$, 则对每对 $\prec k,\ l \succ$, $t_{kl} = \sum_{ij} t_{ij}(\mu_i\bigotimes\nu_j)(\alpha_k,\beta_l) = 0$.

我们也应该注意到, 这个定理完全等价于在第二章第 4 节末尾对 $\mathrm{Hom}(V,W)$ 的基的讨论 .

7.2 多重线性泛函

所有上面的考虑都可推广到多重线性泛函

$$f : V_1 \times \cdots \times V_n \to \mathbb{R} .$$

我们改变一下记号: 如同我们将传统的 $\prec x, y \succ \in \mathbb{R}^2$ 替换成 $\boldsymbol{x} = \prec x_1, \cdots, x_n \succ \in \mathbb{R}^n$ 那样. 因此我们记 $f(\alpha_1, \cdots, \alpha_n) = f(\boldsymbol{\alpha})$, 其中 $\boldsymbol{\alpha} = \prec \alpha_1, \cdots, \alpha_n \succ \in V_1 \times \cdots \times V_n$, 我们的要求是

$$f(\alpha_1, \cdots, \alpha_n)$$

在对所有 $i \neq j$ 的 α_i 固定不动时是 α_j 的线性泛函. 所有这些泛函的集合是个向量空间, 被称作对偶空间 $V_1^*, \cdots, V_n^*$ 的张量积, 记为 $V_1^* \otimes \cdots \otimes V_n^*$.

像以前那样, 在这些张量积空间和各种 Hom 空间之间存在着自然的同构. 例如,

$$V_1^* \bigotimes \cdots \bigotimes V_*^n \quad 与 \quad \mathrm{Hom}(V_1, V_2^* \bigotimes \cdots \bigotimes V_n^*)$$

为自然同构. 但也增加了一类在双线性情形没有遇到过的同构. 对我们来说没有必要深入到这些问题中 .

像前面一样定义*初等*多重线性泛函. 设 $\lambda_i \in V_i^*$, $i = 1, \cdots, n, \boldsymbol{\xi} = \prec \xi_1, \cdots, \xi_n \succ$, 则

$$(\lambda_1 \bigotimes \cdots \bigotimes \lambda_n)(\boldsymbol{\xi}) = \lambda_1(\xi_1) \cdots \lambda(\xi_n) .$$

为使我们的记号能保持尽量简单, 也因为这是我们最感兴趣的情形, 我们只考虑在 $V_1 = V_2 = \cdots = V_n = V$ 的基问题. 这时称 $(V^*)^{\textcircled{n}} = V^* \otimes \cdots \otimes V^*$($n$ 为因子) 为 V 的 n 阶协变张量空间 .

如果 $\{\alpha_j\}_1^m$ 是 V 的基, 且 $f \in (V^*)^{\textcircled{n}}$, 我们可以用关于向量 ξ_i 的基展开式来展开函数值 $f(\boldsymbol{\xi}) = f(\xi_1, \cdots, \xi_n)$, 如同当 f 为双线性情形我们曾做过的那样做, 然而现在的这个结果在记号上更加复杂. 如果我们令 $\xi_i = \sum_{j=1}^m x_j^i \alpha_j$, $i = 1, \cdots, n$(故而 ξ_i 的坐标集为 $\boldsymbol{x}^i = \{x_j^i\}_j$) 并利用 $f(\xi_1, \cdots, \xi_n)$ 对每次一个的每个单独变量的线性性, 我们得到

$$f(\xi_1, \cdots, \xi_n) = \sum x_{p_1}^1 x_{p_2}^2 \cdots x_{p_n}^n f(\alpha_{p_1}, \alpha_{p_2}, \cdots, \alpha_{p_n}) ,$$

其中对所有 n 联组 $\boldsymbol{p} = \prec p_1, \cdots, p_n \succ$ 取和, 使对每个从 1 到 m 的 i 有 $1 \leqslant p_i \leqslant m$. 所有这些 n 联组恰好是从 $\{1, \cdots, n\}$ 到 $\{1, \cdots, m\}$ 的所有函数的集合. 我们

已经把此集合记为 $\overline{m}^{\overline{n}}$, 其中记号 $\overline{n}=\{1,\cdots,n\}$, 从而上面取和的范围可以以下面公式表示：

$$f(\xi_1,\cdots,\xi_n)=\sum_{p\in\overline{m}^{\overline{n}}}x^1_{p_1}\cdots x^n_{p_n}f(\alpha_{p_1},\cdots,\alpha_{p_n})\,.$$

此公式的严格证明可对 n 进行归纳, 留给感兴趣的读者去做吧. 在归纳步骤上他必须将二重求和 $\sum_{p\in\overline{m}^{\overline{n}}}\sum_{j\in\overline{m}}$ 重写为单次求和 $\sum_{q\in\overline{m}^{\overline{(n+1)}}}$, 所用的事实是, 在 $\overline{m}^{\overline{n}}\times\overline{m}$ 中一个有序的偶对 $\prec \boldsymbol{p},j\succ$ 等价于一个 $(n+1)$ 联组 $\boldsymbol{q}\in\overline{m}^{\overline{n+1}}$, 其中对 $i=1,\cdots,n,q_i=p_i$, 而 $q_{n+1}=j$.

如果 $\{\mu_i\}_1^m$ 是 V^* 的对偶基, $\boldsymbol{q}\in\overline{m}^{\overline{n}}$, 设 $\mu_{\boldsymbol{q}}$ 是初等泛函 $\mu_{q_1}\bigotimes\cdots\bigotimes\mu_{q_n}$. 于是, 除了 $\boldsymbol{p}=\boldsymbol{q}$ 外 $\mu_{\boldsymbol{q}}(\alpha_{p_1},\cdots,\alpha_{p_n})=\prod_1^n\mu_{q_i}(\alpha_{p_i})=0$, 而当 $\boldsymbol{p}=\boldsymbol{q}$ 时其值为 1. 更一般地，

$$\mu_{\boldsymbol{q}}(\xi_1,\cdots,\xi_n)=\mu_{q_1}(\xi_1)\cdots\mu_{q_n}(\xi_n)=x^1_{q_1}\cdots x^n_{q_n}\,.$$

因此, 如果我们令 $c_{\boldsymbol{q}}=f(\alpha_{q_1},\cdots,\alpha_{q_n})$, 那么一般展开式的样子是

$$f(\xi_1,\cdots,\xi_n)=\sum_{p\in\overline{m}^{\overline{n}}}c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}(\xi_1,\cdots,\xi_n)\,,$$

或者 $f=\sum c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}$, 它与我们在双线性情形得到的是相同公式, 不过用了更加精致的符号而已. 因此泛函 $\{\mu_{\boldsymbol{p}}:\boldsymbol{p}\in\overline{m}^{\overline{n}}\}$ 张成了 $(V^*)^{ⓝ}$. 它们也是无关的. 因为, 如果 $\sum c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}=0$, 则对每个 $\boldsymbol{q},c_{\boldsymbol{q}}=\sum c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}(\alpha_{q_1},\cdots,\alpha_{q_n})=0$. 我们已经证明了下面的定理 .

定理 2.1 集合 $\{\mu_{\boldsymbol{p}}:p\in\overline{m}^{\overline{n}}\}$ 是 $(V^*)^{ⓝ}$ 的一组基. 对 $(V^*)^{ⓝ}$ 中任意 f, 其坐标函数由 $c_{\boldsymbol{p}}=f(\alpha_{p_1},\cdots,\alpha_{p_n})$ 定义. 因此 $f=\sum c_{\boldsymbol{p}}\mu_{\boldsymbol{p}},f(\xi_1,\cdots,\xi_n)=\sum c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}(\xi_1,\cdots,\xi_n)=\sum c_{\boldsymbol{p}}x^1_{p_1}\cdots x^n_{p_n}$, 对任意 $\prec\xi_1,\xi_2,\cdots,\xi_n\succ\in V^n$ 和任意 $f\in(V^*)^{ⓝ}$ 成立 .

推论 2.1 $(V^*)^{ⓝ}$ 的维数为 m^n .

证明 在 $\overline{m}^{\overline{n}}$ 中有 m^n 个函数, 故基 $\{\mu_{\boldsymbol{p}}:\boldsymbol{p}\in\overline{m}^{\overline{n}}\}$ 有 m^n 个元素 . □

7.3 置换

集合 S 上的一个置换是一个双射 $f:S\to S$. 如果 $\mathcal{S}(S)$ 是 S 上所有置换的集合, 则 $\mathcal{S}=\mathcal{S}(S)$ 在复合下封闭 $(\sigma,p\in\mathcal{S}\Rightarrow\sigma\circ\rho\in\mathcal{S})$ 也在逆运算下封闭

$(\sigma \in \mathcal{S} \Rightarrow \sigma^{-1} \in \mathcal{S})$. 又, 恒同映射 I 也在 $\mathcal{S}$ 中, 当然, 复合运算是结合的. 所有这些论述恰恰说明 $\mathcal{S}$ 在复合运算下是个*群*.除 I 以外的最简单的置换是交换 S 中的一对元素而其他元素保持不动. 称这样的置换为一个*对换*.

现取 S 为有限集 $\overline{n} = \{1, \cdots, n\}$, 令 $\mathcal{S}_n = \mathcal{S}(\overline{n})$. 不难看出, 任一个置换可以表达为对换的乘积, 并且有各种不同的方式 .

一个更为基本的我们所需要的事实是, 如果 ρ 是 $\mathcal{S}_n$ 中一个固定元素, 则映射 $\sigma \mapsto \sigma \circ \rho$ 是 $\mathcal{S}_n \to \mathcal{S}_n$ 的一个双射. 因为任何 σ' 可以写为 $\sigma' = (\sigma' \circ \rho^{-1}) \circ \rho$, 故其为满射的; 又因为 $\sigma_1 \circ \rho = \sigma_2 \circ \rho \Rightarrow (\sigma_1 \circ \rho) \circ \rho^{-1} = (\sigma_2 \circ \rho) \circ \rho^{-1} \Rightarrow \sigma_1 = \sigma_2$, 故其为单射的. 类似地, 映射 $\sigma \mapsto \rho \circ \sigma$($\rho$ 固定) 为双射 .

我们也需要 $\mathcal{S}_n$ 中的元素个数为 $n!$ 这个事实. 这是高中代数所考虑的事. 在定义一个元素 $\sigma \in \mathcal{S}_n$ 时, $\sigma(1)$ 有 n 个选取的方式. 对每个这种选取, $\sigma(2)$ 有 $n-1$ 种选取方式, 从而 $\prec \sigma(1), \sigma(2) \succ$ 可以有 $n(n-1)$ 种选取方式. 对每个这些选取, $\sigma(3)$ 有 $n-2$ 种选取方式, 等等. 总合起来 σ 有 $n(n-1)(n-2)\cdots 1 = n!$ 种选取方式 .

在下文中我们常常写成 '$\rho\sigma$' 来代替 '$\rho \circ \sigma$', 这如同我们有时写成 'ST' 来替代 '$S \circ T$' 这个复合的线性映射一样 .

如果 $\boldsymbol{\xi} = \prec \xi_1, \cdots, \xi_n \succ \in V^n, \sigma \in \mathcal{S}_n$, 我们则可以 "把 σ 施用于 $\boldsymbol{\xi}$" 或 "通过 σ 置换 $\prec \xi_1, \cdots, \xi_n \succ$ 的元素". 我们的意思当然是说我们用 $\prec \xi_{\sigma(1)}, \cdots, \xi_{\sigma(n)} \succ$ 替代了 $\prec \xi_1, \cdots, \xi_n \succ$, 即以 $\boldsymbol{\xi} \circ \sigma$ 替换了 $\boldsymbol{\xi}$.

置换变量把泛函 $f \in (V^*)^{ⓝ}$ 改变为一个新的这种泛函. 特别, 给出 $f \in (V^*)^{ⓝ}$ 和 $\sigma \in \mathcal{S}_n$, 我们定义 f^σ 为

$$f^\sigma(\boldsymbol{\xi}) = f(\boldsymbol{\xi} \circ \sigma^{-1}) = f(\xi_{\sigma^{-1}(1)}, \cdots, \xi_{\sigma^{-1}(n)}) .$$

使用 σ^{-1} 而不是 σ 的部分理由在于它给出了我们下面的公式 .

引理 3.1　$f^{(\sigma_1\sigma_2)} = (f^{\sigma_1})^{\sigma_2}$.

证明　$f^{(\sigma_1\sigma_2)}(\xi) = f(\xi \circ (\sigma_1 \circ \sigma_2)^{-1}) = f(\xi \circ (\sigma_2^{-1} \circ \sigma_1^{-1})) = f((\xi \circ \sigma_2^{-1}) \circ \sigma_1^{-1}) = f^{\sigma_1}(\xi \circ \sigma_2^{-1}) = (f^{\sigma_1})^{\sigma_2}(\xi)$. □

定理 3.1　*对每个 $\sigma \in \mathcal{S}_n$, 由 $f \mapsto f^\sigma$ 定义的映射 T_σ 是由 $(V^*)^{ⓝ}$ 到自己的一个线性同构. 映射 $\sigma \mapsto T_\sigma$ 是由群 $\mathcal{S}_n$ 到 $\mathrm{Hom}((V^*)^{ⓝ})$ 中非异元素的群上的反同态 .*

证明 置换变量不改变多重线性性，故 T_σ 映射 $(V^*)^{ⓝ}$ 到自己. 由于 $(af+bg)^\sigma = af^\sigma + bg^\sigma$，故它是线性的. 并且因为 $f^{\rho\sigma} = (f^\rho)^\sigma$，故 $T_{\rho\sigma} = T_\sigma \circ T_\rho$. 因此 $\sigma \mapsto T_\sigma$ 保持乘积不变但顺序相反. 这就是为什么称它为反同态的理由. 最后，

$$T_{(\sigma^{-1})} \circ T_\sigma = T_{(\sigma\sigma^{-1})} = T_I = I \ ,$$

故 T_σ 可逆 (或非异，或同构) . □

映射 $\sigma \mapsto T_\sigma$ 是群 S_n 用 $(V^*)^{ⓝ}$ 上线性变换的一个表示(其实是一个反表示) .

引理 3.2 *每个 T_σ 将基 $\{\mu_{\boldsymbol{p}}\}$ 带到自己，故是在基上的一个置换* .

证明 我们有 $(\mu_{\boldsymbol{p}})^\sigma(\boldsymbol{\xi}) = \mu_{\boldsymbol{p}}(\boldsymbol{\xi} \circ \sigma^{-1}) = \prod_{i=1}^n \mu_{p_i}(\xi_{\sigma^{-1}(i)})$. 令 $j = \sigma^{-1}(i)$，故有 $i = \sigma(j)$，那么此乘积可被重写为 $\prod_{j=1}^n \mu_{p_{\sigma(j)}}(\xi_j) = \mu_{\boldsymbol{p}\circ\sigma}(\xi)$. 因此

$$(\mu_{\boldsymbol{p}})^\sigma = \mu_{\boldsymbol{p}\circ\sigma} \ ,$$

由于 $\boldsymbol{p} \mapsto \boldsymbol{p} \circ \sigma$ 是 $\overline{m}^{\overline{n}}$ 上的一个置换，故得证 . □

7.4 置换的符号

现在考虑 $\mathbb{R}^n$ 上的特殊的多项式 E，定义为

$$E(\boldsymbol{x}) = E(x_1, \cdots, x_n) = \prod_{1 \leqslant i < j \leqslant n} (x_i - x_j) \ .$$

这是在所有偶对 $\prec i, j \succ \in \overline{n} \times \overline{n}$ 上取的积，其中 $i < j$. 这个有序偶对的集合一一对应于所有二元集 $\{i, j\} \subset \overline{n}, i \neq j$ 的汇集 p，有序偶对由无序偶对按其自然顺序排列. 现在可清楚看到，对任意置换 $\sigma \in S_n$，映射 $\{i, j\} \mapsto \{\sigma(i), \sigma(j)\}$ 是 p 的一个置换. 这意味着在多项式 $E^\sigma(\boldsymbol{x}) = E(\boldsymbol{x}\circ\sigma^{-1})$ 中的因式除了在 σ 倒转了一个偶对的顺序的情形有符号的改变外，与多项式 $E(x)$ 的因式完全一样. 因此，如果 n 是这些倒转的个数，我们则有 $E^\sigma = (-1)^n E$. 映射 $\sigma \mapsto (-1)^n$ 以 'sgn' 表示 (读作 '符号'). 因而 sgn 是从 S_n 到 $\{1, -1\}$ 的一个函数，使得 $E^\sigma = (\text{sgn}\ \sigma)E$，对所有的 $\sigma \in S_n$ 成立. 因为 $(\text{sgn}\ \rho\sigma)E = E^{\rho\sigma} = (E^\rho)^\sigma = (\text{sgn}\ \sigma)E^\rho = (\text{sgn}\ \rho)(\text{sgn}\ \sigma)E$，并可在任意使 $E(\boldsymbol{x}) \neq 0$ 的 n 联组 $\boldsymbol{x}$ 取值 E; 消去因子 $E(\boldsymbol{x})$，得到

$$\text{sgn}\ \rho\sigma = (\text{sgn}\ \rho)(\text{sgn}\ \sigma) \ .$$

如果 σ 是个对换, 我们有

$$\operatorname{sgn} \sigma = -1 .$$

这对 σ 交换相邻的两个数时是很清楚的, 因为这时它正好只改变了 $E(\boldsymbol{x})$ 中一个因子的符号; 一般的情形作为习题留给感兴趣的读者 .

7.5 交错张量子空间 $\mathfrak{a}^n$

定义 5.1 *如果协变张量 $f \in (V^*)^{ⓝ}$ 对所有 $\sigma \in \mathcal{S}_n$ 满足 $f^\sigma = f$ 则称 f 是对称的 .*

如果 f 是双线性的 $[f \in (V^*)^{②}]$, 这正好是条件 $f(\xi,\eta) = f(\eta,\xi)$ 对所有 $\xi,\eta \in V$ 成立 .

定义 5.2 *如果协变张量 $f \in (V^*)^{ⓝ}$ 对所有 $\sigma \in \mathcal{S}_n$ 满足 $f^\sigma = (\operatorname{sgn} \sigma)f$, 则称 f 是反对称的, 或交错的 .*

由于每个 σ 是对换的乘积, 这个定义也可表达成如果它的变量中的两个进行交换时 f 正好改变一次符号这个事实. 在双线性泛函情形就是对所有 $\xi,\eta \in V$ 有 $f(\xi,\eta) = -f(\eta,\xi)$ 这个条件. 有一个重要的情形需要留意, 即只要 n 联组 $\boldsymbol{\xi} = \prec \xi_1,\cdots,\xi_n \succ$ 不是单射的 (即对某些 $i \neq j$ 有 $\xi_i = \xi_j$) 就有 $f(\boldsymbol{\xi}) = 0$. $(V^*)^{ⓝ}$ 中所有对称元素的集合显然是一个子空间, (对我们来说) 更重要的是所有交错元素的集合 $\mathfrak{a}^n$ 也是如此. 有一个重要的从 $(V^*)^{ⓝ}$ 到 $\mathfrak{a}^n$ 的线性投影, 我们现在来描述它 .

定理 5.1 *映射 $f \mapsto (1/n!)\sum_{\sigma\in\mathcal{S}_n}(\operatorname{sgn} \sigma)f^\sigma$ 是一个从 $(V^*)^{ⓝ}$ 到 $\mathfrak{a}^n$ 的投影, 记为 Ω .*

证明 首先验证对每个 $f \in (V^*)^{ⓝ}$ 有 $\Omega f \in \mathfrak{a}^n$. 我们有 $(\Omega f)^\rho = (1/n!)\sum_\sigma(\operatorname{sgn} \sigma)f^{\sigma\rho}$. 而 $\operatorname{sgn} \sigma = (\operatorname{sgn} \sigma\rho)\operatorname{sgn} \rho$. 令 $\sigma' = \sigma \circ \rho$, 并记住 $\sigma \mapsto \sigma'$ 是个双射, 因此我们有

$$(\Omega f)^\rho = \frac{(\operatorname{sgn} \rho)}{n!}\sum_{\sigma'}(\operatorname{sgn} \sigma')f^{\sigma'} = (\operatorname{sgn} \rho)(\Omega f) ,$$

故 $\Omega f \in \mathfrak{a}^n$.

如果 f 已经在 $\mathfrak{a}^n$ 中, 则 $f^\sigma = (\operatorname{sgn} \sigma)f$, 及 $\Omega f = (1/n!)\sum_{\sigma\in\mathcal{S}_n} f$. 因为 S_n 有 $n!$ 个元素, 故 $\Omega f = f$. 所以 Ω 是从 $(V^*)^{\otimes n}$ 到 $\mathfrak{a}^n$ 的投影 . □

引理 5.1 $\Omega(f^\rho) = (\text{sgn } \rho)\Omega f$.

证明 对 $\Omega(f^\rho)$ 的公式与 $(\Omega f)^\rho$ 除了 ρ_δ 替换了 $\sigma\rho$ 外是一样的. 故此证明同于上面定理的证明 . □

定理 5.2 *在 m 维向量空间 V 上的交错 n - 线性泛函空间 $\mathfrak{a}^n$ 的维数等于 $\binom{m}{n}$.*

证明 如果 $f \in \mathfrak{a}^n$ 则 $f = \sum_{\boldsymbol{p}} c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}$, 于是由于 $f^\sigma = (\text{sgn } \sigma)f$, 我们有 $\sum_{\boldsymbol{p}} c_{\boldsymbol{p}}\mu_{\boldsymbol{p}\circ\sigma} = \sum_{\boldsymbol{p}}(\text{sgn } \sigma)c_{\boldsymbol{p}}\mu_{\boldsymbol{p}}$, 其中 σ 为 $\mathcal{S}_n$ 中任意元. 令 $\boldsymbol{p}\circ\sigma = \boldsymbol{q}$, 则左端的和式为 $\sum_{\boldsymbol{q}} c_{\boldsymbol{q}\circ\sigma^{-1}}\mu_{\boldsymbol{q}}$, 而且因为这个基展式是惟一的, 我们对所有 $\boldsymbol{p} \in \overline{m}^{\overline{n}}$ 必有 $c_{\boldsymbol{q}\circ\sigma^{-1}} = (\text{sgn } \sigma)c_{\boldsymbol{q}}$, 或 $c_{\boldsymbol{p}} = (\text{sgn } \sigma)c_{\boldsymbol{p}\circ\sigma}$. 反推过去, 我们看出相反地, 这个条件也蕴涵了 $f^\sigma = (\text{sgn } \sigma)f$. 因此 $f \in \mathfrak{a}^n$ 当且仅当它的坐标函数$c_{\boldsymbol{p}}$ 满足等式

$$c_{\boldsymbol{p}} = (\text{sgn } \sigma)c_{\boldsymbol{p}\circ\sigma}$$

对所有 $\boldsymbol{p} \in \overline{m}^{\overline{n}}$ 和所有 $\sigma \in \mathcal{S}_n$ 成立. 从这个等式可推出许多结果. 其一便是, 除非 $\boldsymbol{p}$ 是一一对应 (单射), 否则 $c_{\boldsymbol{p}} = 0$. 因为如果 $p_i = p_j$, 而 σ 是交换 i 和 j 的对换, 则 $\boldsymbol{p}\circ\sigma = \boldsymbol{p}, c_{\boldsymbol{p}} = (\text{sgn } \sigma)c_{\boldsymbol{p}\circ\sigma} = -c_{\boldsymbol{p}}$, 故 $c_{\boldsymbol{p}} = 0$. 如果 $n > m$, 则没有任何 $\boldsymbol{p}$ 是单射的, 因而我们看出这时 $\mathfrak{a}^n$ 中惟一的元素是零泛函. 那么, $n > m \Rightarrow \dim \mathfrak{a}^n = 0$.

现假设 $n \leqslant m$. 对任何单射的 $\boldsymbol{p}$, 集合 $\{\boldsymbol{p}\circ\sigma : \sigma \in \mathcal{S}_n\}$ 由所有具有与 $\boldsymbol{p}$ 相同值域的 (单射的)n 联组构成. 它们一共有 $n!$ 个. 算进整个值域集合正好只有一个 $\boldsymbol{q} = \boldsymbol{p}\circ\sigma$ 是按其自然序排列的, 即满足 $q_1 < q_2 < \cdots q_n$. 我们选取这个惟一的 $\boldsymbol{q}$ 作为所有具有此值域的元素 $\boldsymbol{p}\circ\sigma$ 的代表. 这些典型的 (代表元)$\boldsymbol{q}$ 的集合 C 因而一一对应于 $\overline{m} = \{1, \cdots, m\}$ 中容积 n 的所有 (值域) 子集的集合 .

每个单射的 $\boldsymbol{p} \in \overline{m}^{\overline{n}}$ 可惟一地表达为 $\boldsymbol{p} = \boldsymbol{q}\circ\sigma$, 其中 $\boldsymbol{q} \in C, \sigma \in \mathcal{S}_n$. 因而 $\mathfrak{a}^n$ 中每个 f 是和式 $\sum_{\boldsymbol{q}\in C}\sum_{\sigma\in\mathcal{S}_n} t_{\boldsymbol{q}\circ\sigma}\mu_{\boldsymbol{q}\circ\sigma}$. 由于 $t_{\boldsymbol{q}\circ\sigma} = (\text{sgn } \sigma)t_{\boldsymbol{q}}$, 这个和式可以重写为 $\sum_{\boldsymbol{q}\in C} t_{\boldsymbol{q}}\sum_\sigma(\text{sgn } \sigma)\mu_{\boldsymbol{q}\circ\sigma} = \sum_{\boldsymbol{q}\in C} t_{\boldsymbol{q}}\nu_{\boldsymbol{q}}$, 其中我们已经设 $\nu_{\boldsymbol{q}} = \sum_\sigma(\text{sgn } \sigma)\mu_{\boldsymbol{q}\circ\sigma} = n!\Omega(\mu_{\boldsymbol{q}})$.

我们几乎要证好了. 由于 $\nu_{\boldsymbol{q}}$ 在 Ω 的值域中, 故每个 $\nu_{\boldsymbol{q}}$ 是交错的, 并且我们刚刚发现对每个 $f \in \mathfrak{a}^n$ 成立的展开式

$$f = \sum_{\boldsymbol{q}\in C} t_{\boldsymbol{q}}\nu_{\boldsymbol{q}}$$

表明集合 $\{\nu_q : q \in C\}$ 张成 $\mathfrak{a}^n$. 它也是无关的, 理由是 $\sum_{q\in C} t_q\nu_q = \sum_{p\in \overline{m}^{\overline{n}}} t_p\mu_p$ 而集合 $\{\mu_p\}$ 是无关的. 因此它是 $\mathfrak{a}^n$ 的一个基.

现在从 $\overline{n} = \{1,\cdots,n\}$ 到 $\overline{m} = \{1,\cdots,m\}$ 的单映射 p 的总的个数是 $m(m-1)\cdots(m-n+1)$, 这是因为第一个元素有 m 种选取方式, 第二个有 $m-1$ 种, 等等直至经 n 次选取, 最后一个元素具有 $m-(n-1)=m-n+1$ 种可能性. 而我们前面已知具有给定值域的这些 p 的个数为 $n!$, 因此不同值域集合的个数为 $m(m-1)\cdots\frac{m-n+1}{n!} = \frac{m!}{n!(m-n)!} = \binom{m}{n}$. 这就是元素 $q \in C$ 的个数. □

$n = m$ 的情形非常重要. 现在 C 只包含一个元素, 即 $\mathcal{S}_m$ 中的单位元 I, 故

$$f = C_I\nu_I = C_I\sum_\sigma(\text{sgn }\sigma)\mu_\sigma$$

从而

$$\begin{aligned}f(\xi_1,\cdots,\xi_m) &= C_I\sum_\sigma(\text{sgn }\sigma)\mu_{\sigma(1)}(\xi_1)\cdots\mu_{\sigma(m)}(\xi_m)\\ &= C_I\sum_\sigma(\text{sgn }\sigma)x^1_{\sigma(1)}\cdots x^m_{\sigma(m)}\ .\end{aligned}$$

我们将看到, 这实质上就是行列式的公式.

7.6 行列式

在第 5 节中我们已经看到在 m 维向量空间 V 上的交错 m 形式的空间 $\mathfrak{a}^n$ 具有维数 $\binom{m}{m}=1$. 因此在纯量因子范围内在 $V=\mathbb{R}^m$ 上有一个惟一的交错的 m 重线性泛函 D, 从而我们可以调整此常数, 使得 $D(\delta^1,\cdots,\delta^m)=1$. 这个惟一确定的 m 形式就是行列式泛函, 其在 m 联组 $\prec \boldsymbol{x}^1,\cdots,\boldsymbol{x}^m \succ$ 的值 $D(\boldsymbol{x}^1,\cdots,\boldsymbol{x}^m)$ 是矩阵$\boldsymbol{x}=(x_{ij})$ 的行列式, 其第 j 列是向量 $\boldsymbol{x}^j, j=1,\cdots,m$.

引理 6.1 $D(\boldsymbol{t}^1,\cdots,\boldsymbol{t}^m) = \sum_{\sigma\in\mathcal{S}_m}(\text{sgn }\sigma)t_{\sigma(1),1}\cdots t_{\sigma(m),m}$.

证明 这恰是上面最后一节的最后面的评注, 这时常数 $C_I=1$(因为 $D(\delta^1,\cdots,\delta^m)=1$), 并将记号改变成了通常的矩阵形式 t_{ij}. □

推论 6.1 $D(\boldsymbol{t}^*) = D(\boldsymbol{t})$.

证明 如果我们按值 σ_i 的顺序去重新排列乘积 $t_{\sigma_1,1}\cdots t_{\sigma_m,m}$ 的因子的顺序, 则此乘积成了 $t_{1,\rho_1}\cdots t_{m,\rho_m}$, 其中 $\rho=\sigma^{-1}$. 由于 $\sigma\mapsto\rho=\sigma^{-1}$ 是从 $\mathcal{S}_m$ 到 $\mathcal{S}_m$ 的

双射, 并且由于 $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}\sigma$, 所以引理中的和式可重写为 $\sum_{\rho\in\mathcal{S}_m}(\operatorname{sgn}\rho)t_{1,\rho_1}\cdots t_{m,\rho_m}$. 但这就是

$$\sum(\operatorname{sgn}\rho)t^*_{\rho_1,1}\cdots t^*_{\rho_m,m} = D(t^*)\ . \qquad \square$$

推论 6.2 *$D(t)$ 是 t 的行向量的交错 m 重线性泛函 .*

现设 $\dim V = m$, 并设 f 为 V 上任意非零交错 m 形式. 对 Hom V 中任意 T, 由 $f_T(\xi_1,\cdots,\xi_n) = f(T\xi_1,\cdots,T\xi_n)$ 定义的泛函 f_T 也属于 $\mathfrak{a}^m$. 由于 $\mathfrak{a}^m$ 是一维的, 对有某个常数 k_T 使 $f_T = k_T f$. 另外, k_T 与 f 无关, 因为, 如果 $g_T = k'_T g$, 而 $g = cf$, 我们必有 $cf_T = k'_T cf$, 从而 $k'_T = k_T$. 称这个惟一决定的常数为 T 的*行列式*; 以 $\Delta(T)$ 表示它. 注意, $\Delta(T)$ 与 V 的任何基的选取无关 .

定理 6.1 $\Delta(S\circ T) = \Delta(S)\Delta(T)$.

证明

$$\begin{aligned}\Delta(S\circ T)f(\xi_1,\cdots,\xi_m) &= f((S\circ T)(\xi_1),\cdots,(S\circ T)(\xi_m))\\ &= f(S(T(\xi_1)),\cdots S(T(\xi_m)))\\ &= \Delta(S)f(T(\xi_1),\cdots,T(\xi_m))\\ &= \Delta(T)\Delta(s)f(\xi_1,\cdots,\xi_m).\end{aligned}$$

现消去 f . $\square$

定理 6.2 *如果 θ 是从 V 到 W 的一个同构, 并且如果 $T\in \mathrm{Hom}V$, $S = \theta\circ T\circ\theta^{-1}$, 则 $\Delta(S) = \Delta(T)$.*

证明 如果 f 是 W 上任意非零交错 m 形式, 并且如果我们定义 g 为 $g(\xi_1,\cdots,\xi_n) = f(\theta\xi_1,\cdots,\theta\xi_n)$. 则 g 为 V 上的非零交错 m 形式 .

现 $f(S\circ\theta\xi_1,\cdots,S\circ\theta\xi_n) = \Delta(s)f(\theta\xi_1,\cdots,\theta\xi_n) = \Delta(S)g(\xi_1,\cdots,\xi_n)$, 又有 $f(S\circ\theta\xi_1,\cdots,S\circ\theta\xi_n) = f(\theta\circ T\xi_1,\cdots,\theta\circ T\xi_n) = g(T\xi_1,\cdots,T\xi_n) = \Delta(T)g(\xi_1,\cdots,\xi_n)$, 从而 $\Delta(s)g = \Delta(T)g$ 从而 $\Delta(S) = \Delta(T)$. $\square$

读者会预料到, 我们所引进的行列式的这两个概念是一致的; 我们现在来证明它 .

推论 6.3 *如果 t 是 T 的关于 V 的某个基的矩阵, 则 $D(t) = \Delta(T)$.*

证明 如果 θ 是个坐标同构, 则 $\overline{T}=\theta\circ T\circ\theta^{-1}$ 在 $\mathrm{Hom}\mathbb{R}^m$ 中且 $\Delta(T)=\Delta(\overline{T})$, 这是上面定理的结果. 又 $\boldsymbol{t}$ 的列是 m 联组 $\overline{T}(\delta^1),\cdots,\overline{T}(\delta^m)$. 因此 $D(\boldsymbol{t})=D(\boldsymbol{t}^1,\cdots,\boldsymbol{t}^m)=D(\overline{T}(\delta^1),\cdots,\overline{T}(\delta^m))=\Delta(\overline{T})D(\delta^1,\cdots,\delta^m)=\Delta(\overline{T})$. 合起来得到 $D(\boldsymbol{t})=\Delta(T)$. □

推论 6.4 *如果 $\boldsymbol{s}$ 和 $\boldsymbol{t}$ 都是 $m\times m$ 矩阵, 则 $D(\boldsymbol{s}\cdot\boldsymbol{t})=D(\boldsymbol{s})D(\boldsymbol{t})$.*

证明 $D(\boldsymbol{s}\cdot\boldsymbol{t})=\Delta(S\circ T)=\Delta(S)\Delta(T)=D(\boldsymbol{s})D(\boldsymbol{t})$. □

推论 6.5 *$D(\boldsymbol{t})=0$ 当且仅当 $\boldsymbol{t}$ 为奇异 .*

证明 如果 $\boldsymbol{t}$ 为非异, 则 $\boldsymbol{t}^{-1}$ 存在, 从而 $D(\boldsymbol{t})D(\boldsymbol{t}^{-1})=D(\boldsymbol{t}\boldsymbol{t}^{-1})=D(I)=1$. 特别, $D(\boldsymbol{t})\neq0$. 如果 $\boldsymbol{t}$ 奇异, 那么某列, 譬中 $\boldsymbol{t}_1$, 是其他列的线性组合, $\boldsymbol{t}_1=\sum_2^m c_i\boldsymbol{t}_i$, 而 $D(\boldsymbol{t}_1,\cdots,\boldsymbol{t}_m)=\sum_2^m c_iD(\boldsymbol{t}_1,\boldsymbol{t}_2,\cdots,\boldsymbol{t}_m)=0$; 其理由是此和式中每一项是 D 在具有两个相等元素的 m 联组上取的值, 所以由交错性质它为 0. □

我们还必须证明 Δ 具有在第二章描述过的所有性质. 其中一些是现成的. 我们知道有 $\Delta(S\circ T)=\Delta(S)\Delta(T)$, 及所有一维和二维的性质是平凡的. 那么, 如果 T 交换二维空间中的无关向量 α_1 和 α_2, 则其矩阵在将其作为基时为 $\boldsymbol{t}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$, 故 $\Delta(T)=D(\boldsymbol{t})=-1$.

下面的引理将完成这项工作.

引理 6.2 *在特殊的假定条件 $\boldsymbol{t}^m=\delta^m$ 下考虑 $D(\boldsymbol{t})=D(\boldsymbol{t}^1,\cdots,\boldsymbol{t}^m)$. 如果 $\boldsymbol{s}$ 是从 $m\times m$ 矩阵中消去最后一行和最后一列得到的 $(m-1)\times(m-1)$ 矩阵, 则 $D(\boldsymbol{s})=D(\boldsymbol{t})$.*

证明 这可由考查引理 6.1 的公式推出, 但我们要直接进行论证 .

如果 $\boldsymbol{t}$ 也以 δ^m 作为它的第 j 列, $j\neq m$, 则由交错性质当然有 $D(\boldsymbol{t})=0$. 这意味着如果第 j 列在其第 m 个位置上有改变时 $D(\boldsymbol{t})$ 不变, 因此 $D(\boldsymbol{t})$ 只依赖于 $i\neq m$ 的行中的 t_{ij} 的值. 就是说, $D(\boldsymbol{t})$ 只依赖于 $\boldsymbol{s}$. 现在 $\boldsymbol{t}\mapsto\boldsymbol{s}$ 显然是到 $\mathbb{R}^{\overline{m-1}\times\overline{(m-1)}}$ 的满映射, 并作为 $\boldsymbol{s}$ 的函数 $D(\boldsymbol{t})$ 是交错 $(m-1)$ 重线性的. 因此它是 $\mathbb{R}^{\overline{(m-1)}\times\overline{(m-1)}}$ 上行列式 D 的常数值. 要知道此常数是什么, 我们在

$$\boldsymbol{s}=\prec\delta^1,\cdots,\delta^{m-1}\succ$$

取值. 于是 $D(\boldsymbol{s})=1=D(\boldsymbol{t})$ 在此特殊值上, 故一般地也有 $D(\boldsymbol{s})=D(\boldsymbol{t})$. □

为了掌握剩下的两个性质, 我们考虑一个 $m \times m$ 矩阵 $\boldsymbol{t}$, 它的最后 $m-n$ 列为 $\delta^{n+1}, \cdots, \delta^{m}$, 我们要重复地使用上面的引理. 首先我们有 $D(\boldsymbol{t}) = D((\boldsymbol{t})^{mm})$, 其中 $(\boldsymbol{t})^{mm}$ 是由 $\boldsymbol{t}$ 去掉最后一行和最后一列所得到的 $(m-1) \times (m-1)$ 矩阵. 因为此矩阵以 δ^{m-1} 为其最后一列 (δ^{m-1} 现在是一个 $(m-1)$ 联组), 相同的论证表明它的行列式同于由它按同样方法得到的 $(m-2) \times (m-2)$ 矩阵. 我们如此继续到最后的 δ 列, 因而看出 $D(\boldsymbol{t})$ 是在 $\boldsymbol{t}$ 的右上角的 $n \times n$ 矩阵的行列式. 如果我们以变换的术语去解释它的话, 我们就得到下面的引理：

引理 6.3 *假设 V 是 m 维的, T 属于 $\mathrm{Hom}V$, 其在一个 $(m-n)$ 维子空间 X 上为恒同映射. 设 Y 为 X 的一个补空间, p 为到 Y 的沿 X 的投影. 则 $p \circ (T \restriction Y)$ 可以考虑成 $\mathrm{Hom}Y$ 中一个元素, 且 $\Delta(T) = \Delta_{(Y)}(p \circ (T \restriction Y))$.*

证明 设 $\alpha_1, \cdots, \alpha_n$ 为 Y 的一组基, $\alpha_{n+1}, \cdots, \alpha_m$ 为 X 的一个基. 于是 $\{\alpha_i\}_1^m$ 为 V 的基, 并且对 $i = n+1, \cdots, m, T(\alpha_i) = \alpha_i$, 那么 T 的矩阵对 $i = n+1, \cdots, m$ 的第 i 列为 δ^i. 如果我们能证明 $\rho \circ (T \restriction Y) \in \mathrm{Hom}Y$ 是 $\boldsymbol{t}$ 的左上角矩阵, 那么引理可由我们上面的讨论得到. 如果同学们能形象地看出向量 $(p \circ T)(\alpha_i), i \leqslant n$ 是什么样子, 那么他们便明白了这点． □

推论 6.6 *在上面情形中, 如果 Y 在 T 下也不变, 则 $\Delta(T) = \Delta_Y(T \restriction Y)$．*

证明 由于现在 $p \circ (T \restriction Y) = T \restriction Y$, 故立即得到证明． □

如果 X 和 Y 的角色互换, 并都在 T 下不变, 而 T 在 Y 上为恒同映射, 则同一引理告诉我们有 $\Delta(T) = \Delta_X(T \restriction X)$．如果我们只知道 X 和 Y 为 T - 不变, 则我们可以分解 T 为一个可交换的乘积 $T = T_1 \circ T_2 = T_2 \circ T_1$, 其中 T_1 和 T_2 是上面所讨论的两个较为特殊的类型, 故有规则 $\Delta(T) = \Delta(T_1)\Delta(T_2) = \Delta_X(T \restriction X)\Delta_Y(T \restriction Y)$, 这是第二章中所列出的另一个性质．

最后的一个规则也是上面引理的一个推论. 如果 T 在 X 上为恒同映射并同时在 V/X 上也为恒同, 则不难看出 $p \circ (T \restriction Y)$ 作为 $\mathrm{Hom}Y$ 中的元素也是恒同的, 故由引理知 $\Delta(T) = 1$．

我们现在证明有关 "按子式 (或余因子)展开" 的定理. 设 $\boldsymbol{t}$ 为 $m \times m$ 矩阵, 并设 $(\boldsymbol{t})^{pr}$ 为由 t 消去第 p 行和第 r 列后得到的 $(m-1) \times (m-1)$ 子矩阵. 于是有

定理 6.3 *$D(\boldsymbol{t}) = \sum_{i=1}^{m} (-1)^{i+r} t_{ir} D((\boldsymbol{t})^{ir})$. 就是说, 我们可以按第 r 列展开*

来对 $D(t)$ 计值; 把第 r 列中每个元素乘上与它相伴的 $(m-1)\times(m-1)$ 矩阵的行列式而后相加. 定理中出现的两个 'D' 自然各具维数 m 和 $m-1$.

证明　在 $t^r=\delta^p$ 的特殊假定下考虑 $D(t)=D(t^1,\cdots,t^m)$. 因为 $D(t)$ 既对 t 的列也对 t 的行均是交错线性泛函, 故而我们可以把第 r 列和第 p 行移到右边缘和底边缘上, 然后用引理 6.2. 因此

$$D(t)=(-1)^{m-r}(-1)^{m-p}D(t^{pr})=(-1)^{p+r}D(t^{pr})\ .$$

一般情形下, $t^r=\sum_{i=1}^m t_{ir}\delta^i$, 如果我们对这个在第 r 位置的和式展开 $D(t^1,\cdots,t^m)$, 另外如果我们用上面对所得和的各分别项的计值, 便得到 $D(t)=\sum_{i=1}^n(-1)^{i+r}t_{ir}\ D((t)^{ir})$. □

推论 6.7　如果 $s\neq r$, 则 $\sum_{i=1}^m(-1)^{i+r}t_{is}D((t)^{ir})=0$.

证明　现在我们有矩阵第 s 列与第 r 列完全相同时的定理的展开式, 由交错性, 这个矩阵的行列式为 0 . □

为了使记号更简明, 令 $c_{ij}=(-1)^{i+j}D((t)^{ij})$. 称其为元素 t_{ij} 在 t 中的余因子. 我们的两个结果合起来是说

$$\sum_{i=1}^m c_{ir}t_{is}=\delta_r^s D(t)\ .$$

特别, 如果 $D(t)\neq 0$, 则分量为 $s_{ri}=c_{ir}/D(t)$ 的矩阵 s 是 t 的逆. 这种观察到的性质给了我们表达一组线性方程解的干净利落的方法. 我们要以 y 解出 $t\cdot x=y$ 的 x, 假定 $D(t)\neq 0$. 于是因为 s 是 t 的逆, 便有 $x=s\cdot y$. 即对 $j=1,\cdots,m, x_j=\sum_{i=1}^m s_{ji}y_i=(\sum_{i=1}^m y_ic_{ij})/D(t)$. 按照我们的展开定理, 此表达式的分子正是由从 t 中以 m 联组 y 替换其第 j 列得到的矩阵的行列式 d_j. 因此, 以这样定义的 d_j, $t\cdot x=y$ 的解是 m 联组

$$x=\prec\frac{d_1}{D(t)},\cdots,\frac{d_m}{D(t)}\succ\ .$$

这是克拉默法则.这与 2.5 节的陈述在记号上稍有不同 .

7.7　外代数

我们最后的工作是在交错 n 重线性泛函 (也称为外 n - 形式)之间引进乘法运算. 我们首先扩张张量乘积运算, 我们需用此运算由泛函形成了初等协变张

量.

定义 7.1 如果 $f \in (V^*)^{ⓝ}$, $g \in (V^*)^{ⓛ}$, 则 $f \bigotimes g$ 是 $(V^*)^{ⓝ+ⓛ}$ 中的元素, 其定义为

$$f \bigotimes g(\xi_1, \cdots, \xi_{n+l}) = f(\xi_1, \cdots, \xi_n) g(\xi_{n+1}, \cdots, \xi_{n+l}) .$$

我们自然要问, 如何将此运算与 $(V^*)^{ⓝ+ⓛ}$ 到 $\mathfrak{a}^{n+l}$ 的投影 Ω 结合起来 .

定理 7.1 $\Omega(f \bigotimes g) = \Omega(f \bigotimes \Omega g) = \Omega(\Omega f \bigotimes g)$.

证明 我们有

$$\begin{aligned} \Omega(f \bigotimes \Omega g) &= \frac{1}{(n+l)!} \sum_{\sigma} (\text{sgn } \sigma)(f \bigotimes \Omega g)^{\sigma} \\ &= \frac{1}{(n+l)!} \sum_{\sigma} (\text{sgn } \sigma) \left(f \bigotimes \frac{1}{l!} \sum_{\rho} (\text{sgn } \rho) g^{\rho} \right)^{\sigma} \\ &= \frac{1}{(n+l)!\, l!} \sum_{\sigma, \rho} (\text{sgn } \sigma)(\text{sgn } \rho)(f \bigotimes g^{\rho})^{\sigma} . \end{aligned}$$

我们可以把 ρ 看作作用在 $f \bigotimes g$ 的全部 $n+l$ 位置上的置换, 但在前 n 个位置为恒同作用. 因而 $(f \bigotimes g^{\rho})^{\sigma} = (f \bigotimes g)^{\rho\sigma}$. 令 $\rho\sigma = \sigma'$. 对每个 σ', 满足 $\rho\sigma = \sigma'$ 的偶对 $\prec \rho, \sigma \succ$ 准确地有 $l!$ 个, 即偶对 $\{\prec \rho, \rho^{-1}\sigma' \succ : \rho \in \mathcal{S}_l\}$. 因此上面的和式为

$$\frac{1}{(n+l)!} \sum_{\sigma'} (\text{sgn } \sigma')(f \bigotimes g)^{\sigma'} = \Omega(f \bigotimes g) .$$

对 $\Omega(\Omega f \bigotimes g)$ 的证明本质上是一样的 . □

定义 7.2 如果 $f \in \mathfrak{a}^n, g \in \mathfrak{a}^l$, 则 $f \wedge g = \binom{n+l}{n} \Omega(f \bigotimes g)$.

引理 7.1 $f_1 \wedge f_2 \wedge \cdots \wedge f_k = (n!/n_1!n_2! \cdots n_k!)\Omega(f_1 \bigotimes \cdots \bigotimes f_k)$, 其中 n_i 是 f_i 的阶, $i = 1, \cdots, k$, 以及 $n = \sum_1^k n_i$.

证明 用楔运算 $\wedge$ 的定义和上面定理直接用归纳法证明. □

推论 7.1 如果 $\lambda_i \in V^*, i = 1, \cdots, n$, 则

$$\lambda_1 \wedge \cdots \wedge \lambda_n = n!\Omega(\lambda_1 \bigotimes \cdots \bigotimes \lambda_n) .$$

特别, 如果 $q_1 < \cdots < q_n$, 且 $\{\mu_i\}_1^m$ 是 V^* 的一个基, 则

$$\mu_{q_1} \wedge \cdots \wedge \mu_{q_n} = n!\Omega(\mu_q) = \mathfrak{a}^n\text{的基元 } \nu_q .$$

定理 7.2 如果 $f \in \mathfrak{a}^n$ 和 $g \in \mathfrak{a}^l$, 则 $g \wedge f = (-1)^{ln} f \wedge g$. 特别对 $\lambda \in V^*$ 有 $\lambda \wedge \lambda = 0$.

证明 如果 σ 是将最后面的 l 个位置中的每一个移动过最前面 n 个位置中的每个元, 我们则有 $g \bigotimes f = (f \bigotimes g)^\sigma$. 因此 σ 是 ln 个交换的积, 从而 $(\operatorname{sgn} \sigma) = (-1)^{ln}$, 并且

$$\Omega(g \bigotimes f) = \Omega(f \bigotimes g)^\sigma = (\operatorname{sgn} \sigma)\Omega(f \bigotimes g) = (-1)^{ln}\Omega(f \bigotimes g) .$$

乘上 $\binom{n+l}{n}$ 便得到定理 . □

推论 7.2 如果 $\{\lambda_i\}_1^n \subset V^*$, 则 $\lambda_1 \wedge \cdots \wedge \lambda_n = 0$ 当且仅当序列 $\{\lambda_i\}_1^n$ 相关.

证明 如果 $\{\lambda_i\}$ 无关, 它便可以扩张为 V^* 的一组基, 从而由上一个推论知 $\lambda_1 \wedge \cdots \wedge \lambda_n$ 是 $\mathfrak{a}^n$ 中的某一个基向量 ν_q, 特别, $\lambda_1 \wedge \cdots \wedge \lambda_n \neq 0$.

如果 $\{\lambda_i\}$ 相关, 则其中一个元素, 譬如 λ_1, 是其余的线性组合, $\lambda_1 = \sum_2^n c_i\lambda_i$, 从而 $\lambda_1 \wedge \lambda_2 \wedge \cdots \wedge \lambda_n = \sum_{i=2}^n c_i\lambda_i \wedge (\lambda_2 \wedge \cdots \wedge \lambda_n)$. 这些项的第 i 个重复出现了 λ_i, 故由引理及上一个推论应为 0 . □

引理 7.2 映射 $\prec f, g \succ \mapsto f \wedge g$ 是由 $\mathfrak{a}^n \times \mathfrak{a}^l$ 到 $\mathfrak{a}^{n+l}$ 的双线性映射 .

证明 由 $f \bigotimes g$ 的显然的双线性性立即推导出 . □

我们以一个重要的扩张定理结束 .

定理 7.3 设 θ 为交错 n 重线性映射

$$\prec \lambda_1, \cdots, \lambda_n \succ \mapsto \lambda_1 \wedge \cdots \wedge \lambda_n ,$$

它从 $(V^*)^n$ 到 $\mathfrak{a}^n$. 于是对任意的交错 n 重线性泛函 $F(\lambda_1, \cdots, \lambda_n) \in (V^*)^n$, 存在 $\mathfrak{a}^n$ 上惟一决定的线性泛函 G, 使得 $F = G \circ \theta$. 因而映射 $G \mapsto F$ 是从 $(\mathfrak{a}^n)^*$ 到 $\mathfrak{a}^n(V^*)$ 的典型同构 .

证明 证明此定理的直接方法是用方程 $F = G \circ \theta$ 确定 G 在一组基上必须的值来定义 G, 然后由 G 的线性性 ,

$$\prec \lambda_1, \cdots, \lambda_n \succ \mapsto \lambda_1 \wedge \cdots \wedge \lambda_n$$

的交错多重线性性和 F 的交错多重线性性表明了等式 $F=G\circ\theta$ 处处成立. 其计算在记号上相当复杂. 我们将用其他方法绕过这种计算. 我们将会看到, 在证明比此定理的结论更多的东西时反倒得到了此定理的一个较短的证明 . □

考虑空间 $\mathfrak{a}^n(V^*)$, 它是 $(V^*)^n$ 上所有交错的 n 重线性函数. 由定理 5.2 知 $d(\mathfrak{a}^n(V^*))=\binom{m}{n}$(因为 $d(V^*)=d(V)=m$). 现对每个 $(\mathfrak{a}^n)^*$ 中泛函 G, 泛函 $G\circ\theta$ 是交错和 n 重线性的, 故 $G\mapsto F=G\circ\theta$ 是由 $(\mathfrak{a}^n)^*$ 到 $\mathfrak{a}^n(V^*)$ 的映射, 它显然是线性的. 它又是个单射, 这是因为, 如果 $G\neq 0$, 则 $F(\mu_{q(1)},\cdots,\mu_{q(n)})=G(\nu_q)\neq 0$, 其中 $\nu_g=\mu_{q(1)}\wedge\cdots\wedge\mu_{q(n)}$ 是 $\mathfrak{a}^n(V^*)$ 的某个基向量. 因为 $d(\mathfrak{a}^n(V^*))=\binom{m}{n}=d(\mathfrak{a}^n)=\mathrm{d}((\mathfrak{a}^n)^*)$, 故此映射是个同构 (由第二章定理 2.4 的推论). 特别地, 每个在 $\mathfrak{a}^n(V^*)$ 中的 F 具有形式 $G\circ\theta$.

可以进一步证明, 上面定理所断言的性质是 $\mathfrak{a}^n$ 的一个抽象的特征描述. 我们这样说的意思是这样的: 假定给出了一个向量空间 X 和从 $(V^*)^n$ 到 X 的一个交错映射 φ, 并假定 $(V^*)^n$ 上每个交错泛函 F 惟一地扩张为 X 上的线性泛函 G(即 $F=G\circ\varphi$). 则 X 同构于 $\mathfrak{a}^n$, 并使 φ 成为 θ .

要明白这点我们只需简单地注意到惟一可扩张性的假定完全准确地等于 $\Phi: G\mapsto F=G\circ\varphi$ 是从 X^* 到 $\mathfrak{a}^n(V^*)$ 的一个同构这个假定. 上面定理给出了一个同构 $\Theta:(\mathfrak{a}^n)^*\to\mathfrak{a}^n(V^*)$ 而伴随映射 $(\Phi^{-1}\circ\Theta)^*$, 因而是由 X^{**} 到 $(\mathfrak{a}^n)^{**}$ 的一个同构, 就是说, 从 X 到 $\mathfrak{a}^n$. 我们将不去验证 φ"成为"θ 了 .

由于定理 6.2 的推论 6.3, 等式 $D(t)=D(t^*)$ 是更加一般的等式 $\Delta(T)=\Delta(T^*)$ 的矩阵形式, 留意这个方程的 "不靠坐标" 的证明是颇有意思的. 当然, 这里的 $T\in\mathrm{Hom}\,V$.

首先注意到等式 $(T^*\lambda)(\xi)=\lambda(T(\xi))$ 通过 $\bigotimes$ 和 $\wedge$ 的定义可以给出

$$T^*\lambda_1\wedge\cdots\wedge T^*\lambda_n(\xi_1,\cdots,\xi_n)=\lambda_1\wedge\cdots\wedge\lambda_n(T\xi_1,\cdots,T\xi_n)\,. \qquad (*)$$

又, $ev_\xi:\lambda_1\wedge\cdots\wedge\lambda_n\mapsto\lambda_1\wedge\cdots\wedge\lambda_n(\xi_1,\cdots,\xi_n)$ 对每个 $\xi\in V^n$ 是一个 $\mathfrak{a}_n(V^*)$ 上的交错 n 重线性泛函. (*) 左端那个项因而是 $ev_\xi(T^*\lambda_1,\cdots,T^*\lambda_n)$; 如果 $n=\dim V$ 由 Δ 的定义, 这就是 $\Delta(T^*)ev_\xi(\lambda_1,\cdots,\lambda_n)$. 由同样的定义, (*) 的右端变成了

$$\Delta(T)[\lambda_1\wedge\cdots\wedge\lambda_n(\xi_1,\cdots,\xi_n)]=\Delta(T)ev_\xi(\lambda_1,\cdots,\lambda_n)\,.$$

因此 (*) 意味着等式 $\Delta(T^*)ev_\xi=\Delta(T)ev_\xi$. 由于当 $\xi=\{\xi_i\}_1^n$ 为无关时, $ev_\xi\neq 0$, 我们便证明了 $\Delta(T^*)=\Delta(T)$.

我们称泛函 $\lambda_i\in V^*$ 的楔积 $\lambda_1\wedge\cdots\wedge\lambda_n$ 为一个*多重向量* .在上面我们看到, $\lambda_1\wedge\cdots\wedge\lambda_n\neq 0$ 当且仅当 $\{\lambda_i\}_1^n$ 是无关的, 这时 $\{\lambda_i\}_1^n$ 张成 V^* 中一个 n 维子

空间. 下面的引理表明这个几何联系并非偶然 .

引理 7.3　V^* 中两个无关的 n 联组 $\{\lambda_i\}_1^n$ 和 $\{\mu_i\}_1^n$ 在 V^* 中张成相同的线性子空间当且仅当 $\mu_1 \wedge \cdots \wedge \mu_n = k(\lambda_1 \wedge \cdots \wedge \lambda_n)$, 其中 k 为某个常数 .

证明　如果 $\{\mu_j\}_1^n \subset L(\{\lambda_i\}_1^n)$, 则每个 μ_j 是这些 λ_i 的线性组合. 如果我们按照这些基的展开式去展开 $\mu_1 \wedge \cdots \wedge \mu_n$. 得到了 $k(\lambda_1 \wedge \cdots \wedge \lambda_n)$. 如果更进一步, $\{\mu_i\}_1^n$ 是无关的, 则 k 非零 .

现设若相反, $\mu_1 \wedge \cdots \wedge \mu_n = k(\lambda_1 \wedge \cdots \wedge \lambda_n), k \neq 0$. 这首先表示 $\{\mu_i\}_1^n$ 是无关的, 并且对每个 j 有

$$\mu_j \wedge (\lambda_1 \wedge \cdots \wedge \lambda_n) = 0 \ ,$$

故每个 μ_j 与 $\{\lambda_i\}_1^n$ 相关. 合起来, 这两个结果蕴涵了集合 $\{\mu_i\}_1^n$ 与 $\{\lambda_i\}_1^n$ 有相同的线性张成空间 . □

这个引理表明一个多重向量有一种与它决定的子空间的关系, 这关系类似于单个向量和它张成空间的关系 .

7.8　内积空间的外幂

设 V 为有限维向量空间, 并设 (,) 为一个非退化 (非异) 的 V 上的对称双线性形式. 在这一节和下一节中我们称任意这种双线性形式为一个内积, 那怕它可能不是正定的. 我们知道这种双线性形式 (,) 诱导了 V 与 V^* 的同构, 它将每个 y 映到 $\overline{y} \in V^*$, 其中 $\overline{y}(x) = \prec x, \overline{y} \succ = (x, y)$ 对所有 x 成立. 于是我们在 V^* 上令 $(\overline{u}, \overline{v}) = (u, v)$ 得到一个非退化的形式 (内积), 并继续用 (,) 表示. 在 $\mathfrak{a}^q$ 上也得到一个非退化内积, 这时令

$$(\overline{u}_1 \wedge \cdots \wedge \overline{u}_q, \overline{v}_1 \wedge \cdots \wedge \overline{v}_q) = \det((\overline{u}_i, \overline{v}_j)). \tag{8.1}$$

要验证 (8.1) 式有意义, 首先注意到对固定的 $\overline{v}_1, \cdots, \overline{v}_q \in V^*$, (8.1) 式的右端是个反对称的向量 $\overline{u}_1, \cdots, \overline{u}_q$ 的多重线性函数, 因此由定理 7.3 可扩张到 $\mathfrak{a}^q(V^*)$ 上. 类似地, 固定这些 $\overline{u}$ 便决定了一个 $\mathfrak{a}^q(V^*)$ 上的线性函数, 从而 (8.1) 式确有定义, 并扩张到 $\mathfrak{a}^q(V^*)$ 上一个双线性函数.(8.1) 式的右端显然对 u 和 v 是对称的, 故此双线性形式确为对称. 要明白它是非退化的, 我们来选取一组基 $u_1, \cdots, u_n$, 使得

$$(u_i, u_j) = \pm\delta_{ij} \ . \tag{8.2}$$

(由第二章定理 7.1, 我们总能找到这样一组基.) 我们知道 $\{\overline{u}_i\} = \{\overline{u}_{i_1} \wedge \cdots \wedge \overline{u}_{i_q}\}$ 构成 $\mathfrak{a}^q$ 的一组基, 其中 $i = \prec i_1, \cdots, i_q \succ$ 遍历所有满足 $1 \leqslant i_1 < \cdots < i_q \leqslant n$ 的整数的 q 联组; 我们断言

$$(\overline{u}_i, \overline{u}_j) = \pm\delta_{ij} . \tag{8.3}$$

事实上, 如果 $i \neq j$, 则对 1 与 q 之间的某个 r 值和对所有 s 有 $i_r \neq j_s$. 这时矩阵 (u_{i_l}, u_{j_m}) 的一整行即第 r 行全都为 0. 因此 (8.1) 式在此时为 0. 如果 $i = j$, 则 (8.2) 式说的是此矩阵沿对角线为 ±1 而其余处为 0, 这就建立了 (8.3) 式, 因此 (,) 在 $\mathfrak{a}^q$ 上非退化. 特别地, 我们有

$$(u_1 \wedge \cdots \wedge u_n, u_1 \wedge \cdots \wedge u_n) = (-1)^{\#} , \tag{8.4}$$

其中 # 是在 (8.3) 式中负号的个数 .

7.9 星号算子

设 V 是个有限维向量空间, 并赋以像第 8 节中那样的非退化内积. 如果 n 是 V 的维数, 则 $\mathfrak{a}^n$ 是一维的. 那么, 在 $\mathfrak{a}^n$ 上所诱导的内积是非退化的, 故 $(\overline{u}, \overline{u})$ 对 $\mathfrak{a}^n$ 中所有非零的 $\overline{u} \in \mathfrak{a}^n$ 要么总是正的, 要么总是负的. 特别地, 在 $\mathfrak{a}^n$ 中正好有两个 $\overline{u}$, 使 $(\overline{u}, \overline{u}) = \pm1$. 让我们选定其中一个, 在本节剩下的内容中认定它不变. 几何地说, 这等于在 V 上选取了一个定向. 我们已经排出了一个

$$(\overline{u}) \in \mathfrak{a}^n, \text{满足} (\overline{u}, \overline{u}) = (-1)^{\#} . \tag{9.1}$$

令 $\overline{v}$ 为 $\mathfrak{a}^q$ 中某个固定的元素. 于是对任意 $\overline{y} \in \mathfrak{a}^{n-q}$ 有 $\overline{v} \wedge \overline{y} \in \mathfrak{a}^n$, 故可写为 $\overline{v} \wedge \overline{y} = f_{\overline{v}}(\overline{y})\overline{u}$, 其中 $f_{\overline{v}}(\overline{y})$ 线性依赖于 $\overline{y}$. 因为在 $\mathfrak{a}^{n-q}$ 上所诱导的内积 (,) 是非退化的, 故存在一个惟一的元素 $*\overline{v} \in \mathfrak{a}^{n-q}$ 使得 $(\overline{y}, *\overline{v}) = f_{\overline{v}}(\overline{y})$. 重复说一下, 我们对每个 $\overline{v} \in \mathfrak{a}^q$, 用令

$$(\overline{y}, *\overline{v})\overline{u} = \overline{v} \wedge \overline{y} \tag{9.2}$$

的方式给定了一个 $*\,\overline{v} \in \mathfrak{a}^{n-q}$. 那么我们已定义了从 $\mathfrak{a}^q$ 到 $\mathfrak{a}^{n-q}$ 的一个映射 $*$. 由 (9.2) 式知此映射是线性的. 设 $u_1, \cdots, u_n$ 为满足 (8.2) 式的 V 的一组基, 也令 $\overline{u} = u_1 \wedge \cdots \wedge u_n$, 并构造 $\mathfrak{a}^q$ 和 $\mathfrak{a}^{n-q}$ 的对应的基. 于是如果在 q 联组 i 中出现的任一 i_l 也出现在 j 则 $\overline{u}_i \wedge \overline{u}_j = 0$. 如果没有任何 i_l 出现在 j 中, 则

$$\overline{u}_i \wedge \overline{u}_j = \varepsilon_{i_1 \cdots i_q j_1 \cdots j_{n-q}} \overline{u} ,$$

其中 $\varepsilon_{\boldsymbol{k}} = \operatorname{sgn} \boldsymbol{k}$.

如果将此与 (9.2) 式和 (8.3) 式比较, 我们看出

$$*\overline{u}_{\boldsymbol{i}} = \pm\varepsilon_{i_1\cdots i_q j_1\cdots j_{n-q}}\overline{u}_{\boldsymbol{j}} \ , \tag{9.3}$$

其中的符号用于在 (8.3) 式中出现的, 即符号为正还是负按照满足 $(\overline{u}_{j_l}, \overline{u}_{j_l}) = -1$ 的 j_l 出现在 $\boldsymbol{j}$ 中的个数是偶数还是奇数而定. 将 $*$ 施用于 (9.3) 式, 我们看到, 对所有 $v \in \mathfrak{a}^q$,

$$** \ \overline{v} = (-1)^{q(n-q)+\#}v \ . \tag{9.4}$$

设 v 和 w 为 $\mathfrak{a}^q$ 中元素. 于是

$$\begin{aligned}(*v, *w)u &= v \wedge *w = (-1)^{q(n-q)} * w \wedge v \\ &= (-1)^{q(n-q)}(** \ w, v)u \ .\end{aligned}$$

如果使用 (9.4) 式, 则有

$$(* \, v, * \, w) = (-1)^{\#}(v, w) \ . \tag{9.5}$$

第八章　积分

8.1　引言

本章将讲述在 n 维欧氏空间 $\mathbb{E}^n$ 中的一个积分理论, 读者所记得的欧氏空间, 简单地说就是 n 维笛卡儿空间连同标准的内积. 我们的主要事情就是对 $\mathbb{E}^n$ 的子集引进一个大小的概念 (二维中的面积, 三维中的体积, $\cdots$). 在进行正式定义之前, 让我们看一看我们想要的大小概念要有什么样的性质. 我们要寻找一个函数 μ, 它对有界子集 $A \subset \mathbb{E}^n$ 派定一个数 $\mu(A)$.

(i) 我们想要 $\mu(A)$ 为非负实数 .

(ii) 如果 $A \subset B$, 我们想有 $\mu(A) \leqslant \mu(B)$.

(iii) 如果 A 和 B 不交 (即 $A \cap B = \varnothing$), 则我们期待有 $\mu(A \cup B) = \mu(A) + \mu(B)$.

(iv) 设 T 为任意一个欧几里得运动 [1]. 对任何集合 A, 设 TA 为所有形如 Tx 的点 $x \in A$ 的集合. 于是我们期待有 $\mu(TA) = \mu(A)$. (因而我们需要 "全等" 集合有相同大小 .)

(v) 我们期待一个 "低维集合"(对此有适当的定义) 的大小为 0. 因此直线中的点, 平面中的曲线, 三维空间中的曲面, 等等, 大小应该全都为零 .

(vi) 由相同的理由, 我们期待开集具有正的大小 .

---

[1] 回忆一个欧几里得运动是 $\mathbb{E}^n$ 中的一个等距映射, 因而可以表示为平移与正交变换的复合 .

在上面的讨论中我们没有仔细阐明我们谈的是什么样的集合. 人们或许有极大的愿望试图对 $\mathbb{E}^n$ 中每个子集都给出一个大小. 但这被证明是不可能的, 理由如下: 设 U 和 V 为 $\mathbb{E}^3$ 中任意两个有界开子集. 可以证明 [2] 我们能找到分解

$$U = \bigcup_{i=1}^{k} U_i \quad \text{和} \quad V = \bigcup_{i=1}^{k} V_i$$

满足 $U_i \cap U_j = \varnothing = V_i \cap V_j, i \neq j$ 及欧几里得运动 T_i 使得 $T_i U_i = V_i$. 换句话说, 我们可以分裂 U 为有限多块, 移动它们然后再组合就得到了 V. 不用说, 集合 U_i 必定看起来非常之坏. 我们立刻就会想到, 如果希望对所有子集 (包括像那些 U_i) 给出一个大小度量我们便不能满足 (ii),(iii),(iv) 和 (vi). 事实上 (iii)(重复 $(k-1)$ 次) 意味着

$$\mu(U) = \sum_{i=1}^{k} \mu(U_i) \ ,$$

而 (iv) 蕴涵了 $\mu(U_i) = \mu(V_i)$. 因此 $\mu(U) = \mu(V)$, 或者说, 任何两个开集的大小相同. 因为任一开集包含有两个不交的开集, 这表明由 (ii) 有 $\mu(U) \geqslant 2\mu(U)$, 故 $\mu(U) = 0$.

因此我们面临着选择. 要么去掉上面 (i)~(vi) 中某些要求, 要么我们不对 $\mathbb{E}^n$ 中每个子集都派定一个大小度量. 由于我们的要求是合理的, 我们宁愿选取第二个可能性. 当然这表明现在除了引进大小概念外, 我们还必须描述那些我们所接受的 “好” 的集合的类 .

我们将以公理化的方式进行, 列出一些对一类子集合和一个函数 μ 的 “合理的” 公理 .

8.2 公理

我们的公理关联到 $\mathbb{E}^n$ 的一个子集类 $\mathfrak{D}$ 和定义在 $\mathfrak{D}$ 上的函数 μ .(即如果 A 是 $\mathbb{E}^n$ 中属于我们类 $\mathfrak{D}$ 的子集, 则 $\mu(A)$ 有定义 .)

I. $\mathfrak{D}$ 是 $\mathbb{E}^n$ 的一些子集合的汇集, 满足 :

$\mathfrak{D}$1. 如果 $A \in \mathfrak{D}, B \in \mathfrak{D}$ 则 $A \cup B \in \mathfrak{D}, A \cap B \in \mathfrak{D}$, 还有 $A - B \in \mathfrak{D}$.

$\mathfrak{D}$2. 如果 $A \in \mathfrak{D}$, T 是个平移, 则 $TA \in \mathfrak{D}$.

[2] S. Banach and A. Tarski, Sur la décomposition des ensembles de pointes en partie respectivement congruentes, *Fund. Math.* 6,244~ 277(1924). R. M. Robinson, On the decomposition of Spheres, *Fund. Math.* 34,246~ 260(1947) .

$\mathfrak{D}3$. 集合 $\square_0^1 = \{x : 0 \leqslant x^i < 1\}$ 属于 $\mathfrak{D}$.

II. 实函数 μ 有下列性质:

$\mu1$. 对所有 $A \in \mathfrak{D}, \mu(A) \geqslant 0$.

$\mu2$. 如果 $A \in \mathfrak{D}, B \in \mathfrak{D}$ 且 $A \cap B = \varnothing$, 则 $\mu(A \cup B) = \mu(A) + \mu(B)$.

$\mu3$. 对任何 $A \in \mathfrak{D}$ 和任何平移 T, 有 $\mu(TA) = \mu(A)$.

$\mu4$. $\mu(\square_0^1) = 1$.

继续进行之前, 对我们的公理按顺序作些评注. 公理 $\mathfrak{D}1$ 让我们可以对 $\mathfrak{D}$ 中元素进行初等的集合论中的运算. 注意在公理 $\mathfrak{D}2$ 和 $\mu3$ 中只允许进行平移, 但是在我们所罗列的想要的性质中有 (iv) 中要求关于对所有欧几里得运动有适当表现. 原因是我们将证明对于 $\mathfrak{D}$ 的 "好" 选择, 这些列出的公理惟一地决定了 μ. 虽然我们只设定了较弱的条件 $\mu3$ 作为公理, 但最终将证明, μ 实际上满足了更强的条件 (iv) .

公理 $\mathfrak{D}3$ 保证了我们的理论不是完全平凡的, 即汇集 $\mathfrak{D}$ 是非空的. 公理 $\mu4$ 具有法化 μ 的作用. 没有它, 任何满足 $\mu1, \mu2$ 和 $\mu3$ 的 μ 都可以乘上一个非负实数, 如此得到的新函数 μ' 仍然会满足我们的公理. 特别地, $\mu4$ 确保了我们没有选择 μ 为对每个 A 指派 0 的平凡函数 .

下面几节的计划是对 $\mathfrak{D}$ 作一些合理的选择并证明对所给定的这个 $\mathfrak{D}$ 存在一个惟一的满足 $\mu1$ 到 $\mu4$ 的 μ .

我们常常用到而不加评论的一个 $\mathfrak{D}, \mu$ - 公理的一个重要的初等推论是

$\mu5$. 如果 $A \subset \cup_1^n A_i$, 其中的所有集合均在 $\mathfrak{D}$ 中, 则 $\mu(A) \leqslant \sum_1^n \mu(A_i)$.

我们开始做的大部分事情是组合式的. 我们首先考虑 (广义的) 矩形, 它是区间的笛卡儿积, 以及用在其中一个点来决定一个分裂矩形的方法, 即像在图 8.1

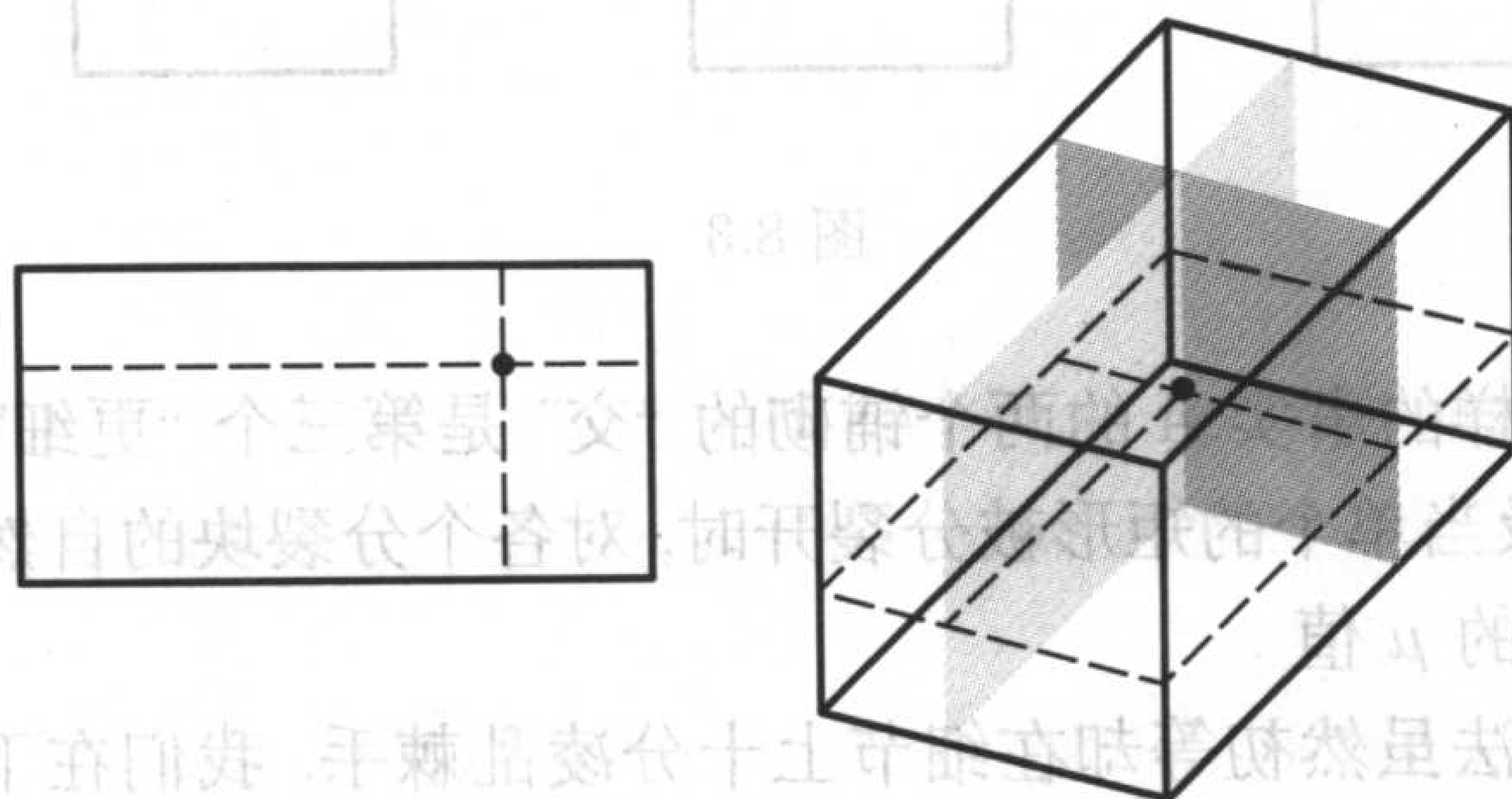


图 8.1

中表示的那样将矩形分成更小的矩形的组合. 它所关联的事实是任意两个矩形的交为矩形, 两个矩形之差为有限个矩形的不交并 (见图 8.2) .

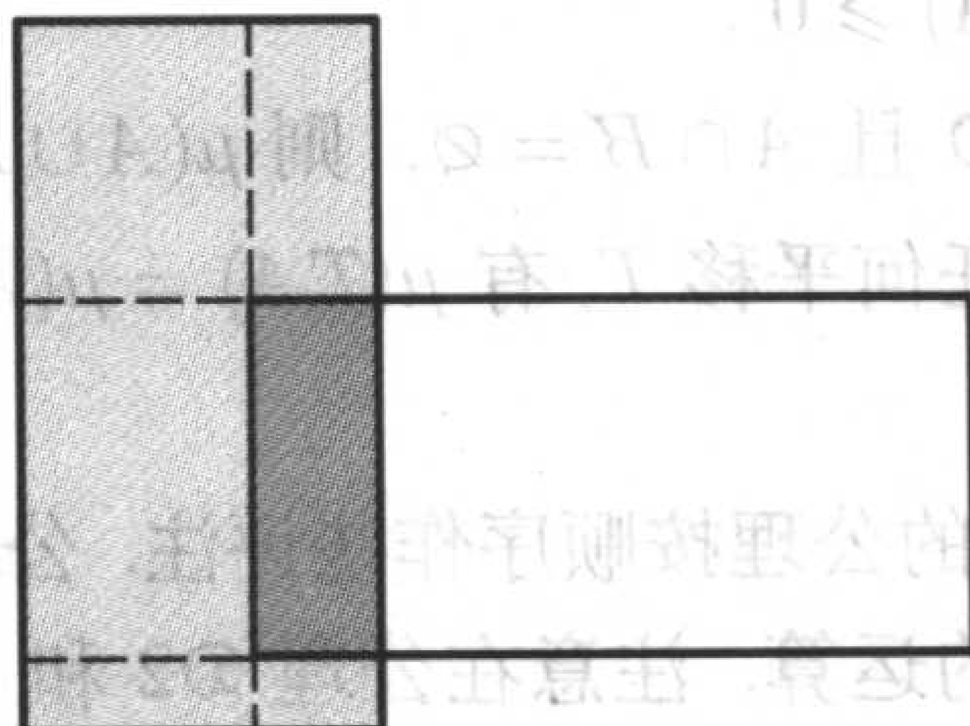


图 8.2

我们称集合 A 是可铺的是说它可以表示为矩形的有限个不交汇集 p 的并集 (称为 A 的一个铺砌). 由我们这种组合式的想法可知, 所有可铺的集合的汇集 $\mathfrak{D}_{\min}$ 满足公理 $\mathfrak{D}1$ 到 $\mathfrak{D}3$, 并且是那样的最小族: 任何其他满足这些公理的汇集 $\mathfrak{D}$ 包含了 $\mathfrak{D}_{\min}$. 于是, 如果 μ 在 $\mathfrak{D}_{\min}$ 上满足 $\mu1$ 到 $\mu4$, 则它必定对一个矩形有自然的值 (即边长的乘积). 这表明, 由于对任意可铺集 $\mu(A)$ 必定是 A 的一个铺砌中矩形自然值的和, 故在 $\mathfrak{D}_{\min}$ 上 μ 由要求 $\mu1$ 到 $\mu4$ 而惟一地决定. 在 $\mathfrak{D}_{\min}$ 上 μ 的存在性依赖于一个关键的引理, 即 A 的两个不同的铺砌给出相同的和数 (见图 8.3) .

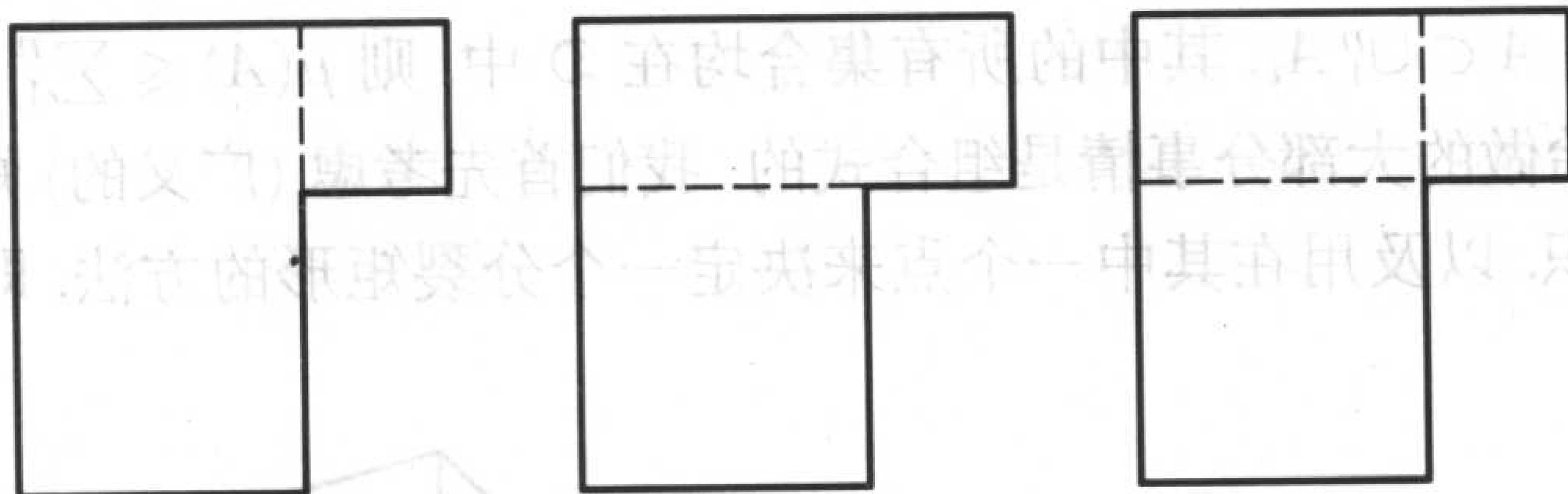


图 8.3

这归结到这样的事实:A 的两个铺砌的 "交" 是第三个 "更细" 于它们中任何一个的铺砌, 以及当单个的矩形被分裂开时, 对各个分裂块的自然值 μ 相加成这个被分段的矩形的 μ 值 .

所有这些想法虽然初等却在细节上十分凌乱棘手. 我们在下面给出的证明是为读者在有疑问时作为参考之用, 他或许只须读一读定义和结果的陈述, 然后便进到第 6 节 .

8.3 矩形和可铺集合

首先引进一些记号和术语. 设 $\boldsymbol{a}=\prec a^1,\cdots,a^n\succ$ 和 $\boldsymbol{b}=\prec b^1,\cdots,b^n\succ$ 为 $\mathbb{E}^n$ 中元素. 我们所说的矩形 $\square_{\boldsymbol{a}}^{\boldsymbol{b}}$ 是指一个集合, 它的所有点 $\boldsymbol{x}=\prec x_1,\cdots,x_n\succ\in\mathbb{E}^n$ 满足 $a^i\leqslant x^i<b^i$. 因此

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}}=\{\boldsymbol{x}:a^i\leqslant x^i<b^i,i=1,\cdots,n\}\,. \tag{3.1}$$

注意, 为了让 $\square_{\boldsymbol{a}}^{\boldsymbol{b}}$ 非空, 必须有 $a^i<b^i$ 对所有 i 成立. 换句话说, 如果 $a^i\geqslant b^i$ 对某个 i 成立, 则

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}}=\varnothing\,. \tag{3.2}$$

例如在平面 $(n=2)$ 中, 我们的矩形对应于普通的欧几里得矩形,其边平行于轴(或许我们应该用一个附加形容词来称呼我们的集合为水平矩形, 支柱形矩形或者别的什么, 但为简明起见, 我们只叫它们做矩形 .) 注意, 在平面中我们的矩形只有左边和下边而没有右边和上边 (见图 8.4) .

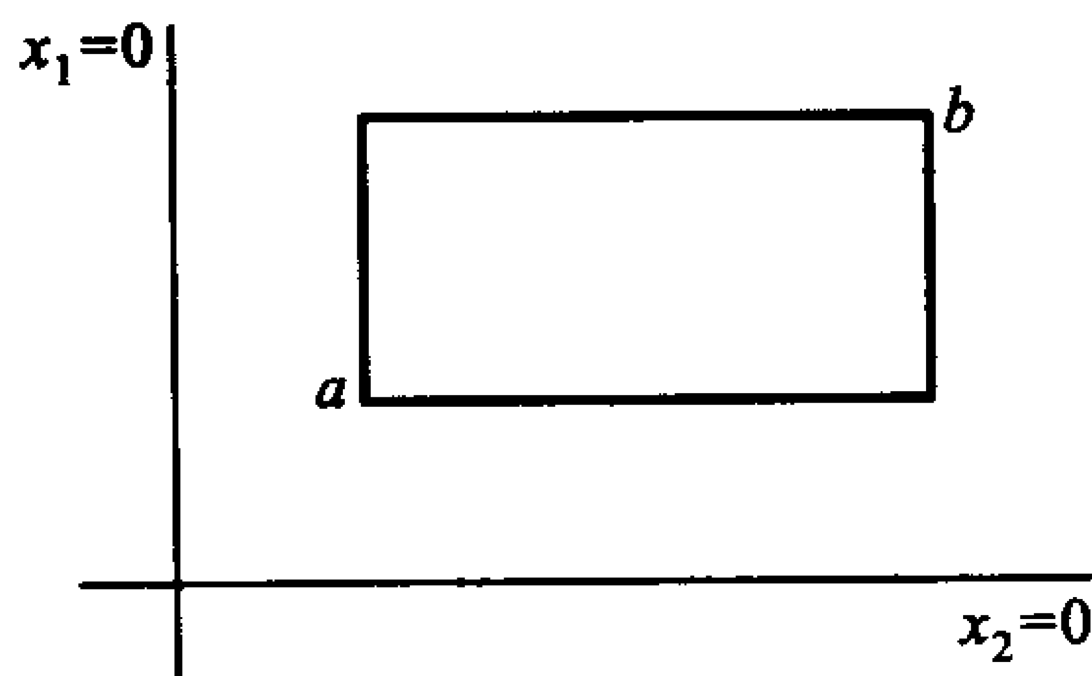


图 8.4

对一般的 n, 如果我们令 $\mathbf{1}=\prec 1,1,\cdots,1\succ$, 那么我们的记号与 §3 的一致.

我们现在汇拢一些关于矩形的初等性质. 从定义 (3.1) 式立即可以推出, 如果 $\boldsymbol{a}=\prec a^1,\cdots,a^k\succ,\boldsymbol{b}=\prec b^1,\cdots,b^k\succ$, 等等, 则

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}}\bigcap\square_{\boldsymbol{c}}^{\boldsymbol{d}}=\square_{\boldsymbol{e}}^{\boldsymbol{f}}, \tag{3.3}$$

其中 $e^i=\max(a^i,c^i),f^i=\min(b^i,d^i),i=1,\cdots,n$. (读者应画出各种不同的在平面上关于这个等式的例子从而得到正确的几何感觉.) 注意此时 $\square_{\boldsymbol{a}}^{\boldsymbol{b}}\cap\square_{\boldsymbol{c}}^{\boldsymbol{d}}=\varnothing$ 的情形由 (3.2) 式也包含在 (3.3) 式中. 另一个由定义 (3.1) 式直接得到的结果是

$$\text{对任意的平移 } T,\ T\square_{\boldsymbol{a}}^{\boldsymbol{b}}=\square_{T\boldsymbol{a}}^{T\boldsymbol{b}}\,. \tag{3.4}$$

我们要建立一些初等的结果, 而它们将意味着对任何满足公理 $\mathfrak{D}1$ 到 $\mathfrak{D}3$ 的 $\mathfrak{D}$ 必包含所有的矩形 .

引理 3.1 任何矩形 $\square_{\boldsymbol{a}}^{\boldsymbol{b}}$ 可以写为不交并

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}}=\bigcup_{r=1}^{k}\square_{\boldsymbol{a}_r}^{\boldsymbol{b}_r},\quad 其中\boldsymbol{b}_r-\boldsymbol{a}_r\in\square_{\boldsymbol{0}}^{\boldsymbol{1}}.$$

(这个引理说的是任何一个 “大的” 矩形可以写为有限个 “小的” 矩形的并.)

证明 我们可假设 $\square_{\boldsymbol{a}}^{\boldsymbol{b}}\neq\varnothing$(否则在并中取 $k=0$). 因此 $b^i>a^i$. 特别, 如果我们选取一个充分大的整数 m, 则 $(1/2^m)(\boldsymbol{b}-\boldsymbol{a})$ 将位于 $\square_{\boldsymbol{0}}^{\boldsymbol{1}}$ 中 .

用归纳法, 只要证明我们能把 $\square_{\boldsymbol{a}}^{\boldsymbol{b}}$ 分解为不交并

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}}=\bigcup_{s=1}^{2^n}\square_{\boldsymbol{c}_s}^{\boldsymbol{d}_s},\qquad 其中\boldsymbol{d}_s-\boldsymbol{c}_s=\frac{1}{2}(\boldsymbol{b}-\boldsymbol{a})\tag{3.5}$$

就可以了.(因为由此我们可以继续细分直到我们得到了足够小的矩形.)

我们以显而易见的方式来得到这个剖分, 即选取矩形 “中间位置的” 点为顶点, 并考虑由通过此点而平行坐标面的超平面分割得到的所有矩形. 要写出一个显式表述的公式, 方便的办法是使用 $\{1,\cdots,n\}$ 中所有子集合的集合为指标集而不用整数 $1,\cdots,2^n$ 为指标. 设 J 表示 $\{1,2,\cdots,n\}$ 的一个任意子集. 设 $\boldsymbol{a}_J=\prec a_J^1,\cdots,a_J^n\succ,\boldsymbol{b}_J=\prec b_J^1,\cdots,b_J^n\succ$ 由

$$a_J^i=\begin{cases}a^i+\frac{b^i-a^i}{2}, & 如果i\in J,\\ a^i, & 如果i\notin J,\end{cases}$$

以及

$$b_J^i=\begin{cases}b^i, & 如果i\in J,\\ b^i-\frac{b^i-a^i}{2}, & 如果i\notin J,\end{cases}$$

给出. 于是任一 $x\in\square_{\boldsymbol{a}}^{\boldsymbol{b}}$ 位于一个且只有一个 $\square_{\boldsymbol{a}_J}^{\boldsymbol{b}_J}$ 中. 换句话说, 如果 $J\neq K$ 则 $\square_{\boldsymbol{a}_J}^{\boldsymbol{b}_J}\cap\square_{\boldsymbol{a}_k}^{\boldsymbol{b}_k}=\varnothing$, 并且 $\cup_{所有J}\square_{\boldsymbol{a}_J}^{\boldsymbol{b}_J}=\square_{\boldsymbol{a}}^{\boldsymbol{b}}$. ($n=2$ 的情形展示在图 8.5 中.) 因为 $\boldsymbol{b}_J-\boldsymbol{a}_J=\frac{1}{2}(\boldsymbol{b}-\boldsymbol{a})$ 对所有 J 成立, 故引理得证 . □

现在我们看到, 由 (3.3) 式对任意 $\boldsymbol{c}\in\square_{\boldsymbol{0}}^{\boldsymbol{1}}$ 我们有

$$\square_{\boldsymbol{0}}^{\boldsymbol{c}}=\square_{\boldsymbol{0}}^{\boldsymbol{1}}\cap\square_{\boldsymbol{c}-\boldsymbol{1}}^{\boldsymbol{c}}.\tag{3.6}$$

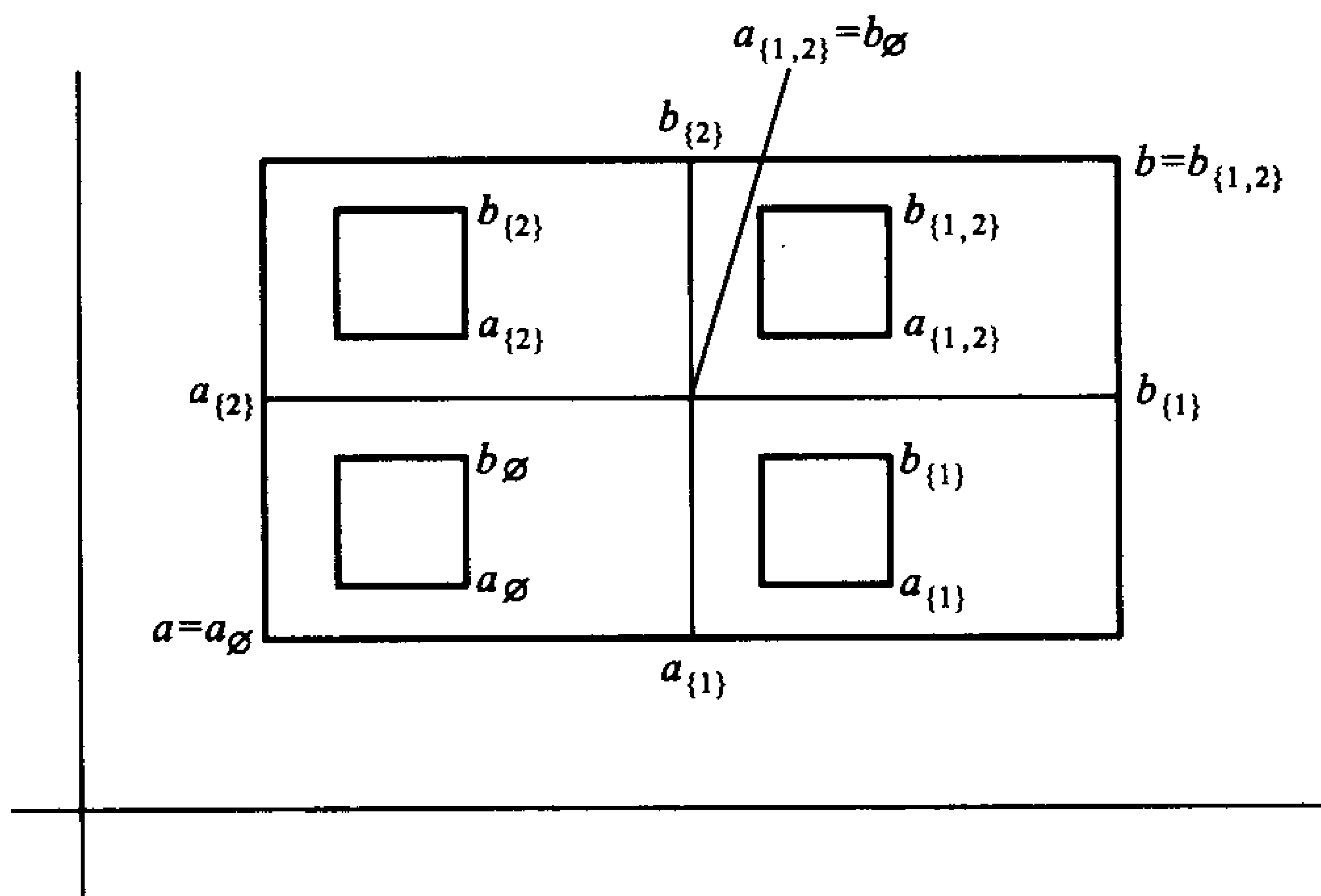


图 8.5

设 T_v 表示通过向量 v 的平移. 于是 $\square_{\boldsymbol{c-1}}^{\boldsymbol{c}} = T_{\boldsymbol{c-1}}\square_{\boldsymbol{0}}^{\boldsymbol{1}}$, 这由 (3.4) 式得出. 因此由公理 $\mathfrak{D}2$ 和 $\mathfrak{D}3$ 矩形 $\square_{\boldsymbol{c-1}}^{\boldsymbol{c}}$ 必定属于 $\mathfrak{D}$. 由 (3.6) 式和公理 $\mathfrak{D}1$ 我们得出结论, $\square_{\boldsymbol{0}}^{\boldsymbol{c}} \in \mathfrak{D}$ 对任何 $\boldsymbol{c} \in \square_{\boldsymbol{0}}^{\boldsymbol{1}}$ 成立.

由 (3.4) 式观察到 $T_{\boldsymbol{a}}\square_{\boldsymbol{0}}^{\boldsymbol{b-a}} = \square_{\boldsymbol{a}}^{\boldsymbol{b}}$. 因此一旦 $\boldsymbol{b}-\boldsymbol{a} \in \square_{\boldsymbol{0}}^{\boldsymbol{1}}$, 则

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}} \in \mathfrak{D} .$$

如果我们此时用引理 3.1 便得出结论说, 对所有 $\boldsymbol{a}$ 和 $\boldsymbol{b}$,

$$\square_{\boldsymbol{a}}^{\boldsymbol{b}} \in \mathfrak{D} . \tag{3.7}$$

我们给出下面的定义 .

定义 3.1 子集 $S \subset \mathbb{E}^n$ 称为一个可铺集是说, 如果 S 是有限多个矩形的不交并.

我们于是可断言 :

命题 3.1 满足公理 $\mathfrak{D}1$ 到 $\mathfrak{D}3$ 的任何一个 $\mathfrak{D}$ 必定包含了所有的可铺集. 设 $\mathfrak{D}_{\min}$ 表示所有矩形有限并集的汇集, 则 $\mathfrak{D}_{\min}$ 满足公理 $\mathfrak{D}1$ 到 $\mathfrak{D}3$.

证明 我们已经证明了此命题的第一部分. 将第二部分作为习题留给读者 .□

8.4 极小理论

我们将看到由公理 $\mu1$ 到 $\mu3$ 决定的 μ 到底能走多远. 事实上, 我们要证明 $\mu(\square_a^b)$ 正是它应该是的那个样子, 即, 如果

$$\boldsymbol{a} = \prec a^1, \cdots, a^n \succ \quad \text{和} \quad \boldsymbol{b} = \prec b^1, \cdots, b^n \succ ,$$

则我们必有

$$\mu(\square_{\boldsymbol{a}}^{\boldsymbol{b}}) = \begin{cases} 0, & \text{如果}\square_{\boldsymbol{a}}^{\boldsymbol{b}} = \varnothing \\ (b_1 - a_1)\cdots(b_n - a_n), & \text{如果}\square_{\boldsymbol{a}}^{\boldsymbol{b}} \neq \varnothing . \end{cases} \tag{4.1}$$

公理 $\mu4$ 说,(4.1) 式对特殊情形 $\boldsymbol{a} = \boldsymbol{0}, \boldsymbol{b} = \boldsymbol{1}$ 成立. 考查引理 3.1 的证明表明 $\square_{\boldsymbol{0}}^{\boldsymbol{1}}$ 可以写做 2^n 个矩形的不交并, 它们都全等于 (通过平移)$\square_{\boldsymbol{0}}^{\boldsymbol{1/2}}$, 其中 $\boldsymbol{\frac{1}{2}} = (\frac{1}{2}, \cdots, \frac{1}{2})$. 那么公理 $\mu2$ 和 $\mu3$ 表示

$$\mu(\square_{\boldsymbol{0}}^{\boldsymbol{1/2}}) = \frac{1}{2^n} .$$

重复这个论断便归纳地证明了如果 $\boldsymbol{\frac{1}{2^r}} = \prec \frac{1}{2^r}, \cdots, \frac{1}{2^r} \succ$ 则

$$\mu(\square_{\boldsymbol{0}}^{\boldsymbol{1/2^r}}) = \frac{1}{2^{nr}} \tag{4.2}$$

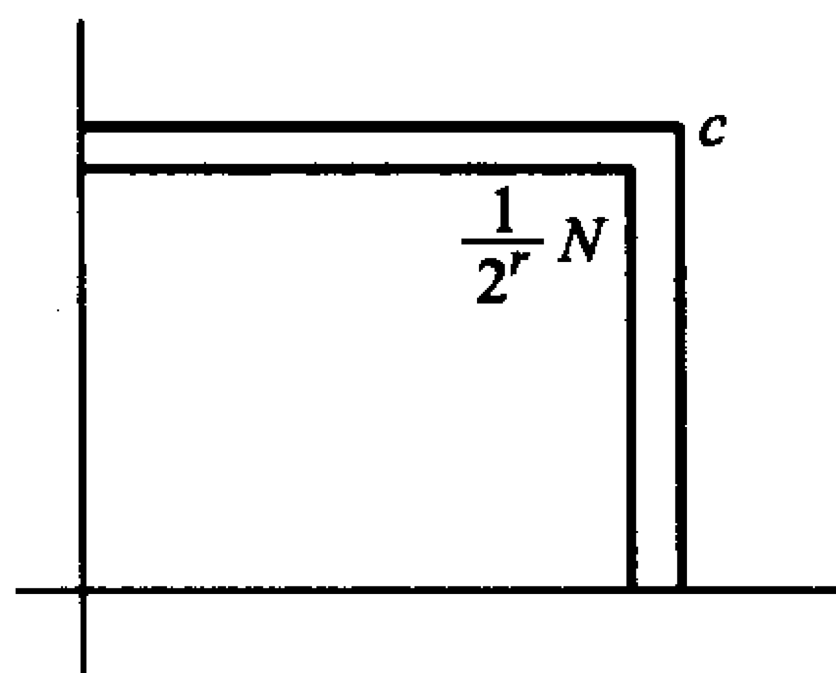


图 8.6

我们要用 (4.2) 式去验证 (4.1) 式. 想法是用立方形

$$\square_{\boldsymbol{0}}^{\boldsymbol{1/2^r}}$$

的平移的并去逼近任一个矩形 . 在证明 (4.1) 式时我们发现, 只需要考虑形如 $\square_{\boldsymbol{0}}^{\boldsymbol{c}}$ 的矩形. 事实上, 取 $\boldsymbol{c} = \boldsymbol{b} - \boldsymbol{a}$, 并看到

$$T_{-\boldsymbol{a}}(\square_{\boldsymbol{a}}^{\boldsymbol{b}}) = \square_{\boldsymbol{0}}^{\boldsymbol{c}} ,$$

故而公理 $\mu3$ 蕴涵了 $\mu(\square_{\boldsymbol{a}}^{\boldsymbol{b}}) = \mu(\square_{\boldsymbol{0}}^{\boldsymbol{c}})$, 并由定义, $c^1, \cdots c^n = (b^1 - a^1) \cdot \cdots \cdot (b^n - a^n)$. 如果 $\square_{\boldsymbol{0}}^{\boldsymbol{c}} = \varnothing$, 则 (4.1) 式显而易见正确 (由公理 $\mu2$) . 假定 $\square_{\boldsymbol{0}}^{\boldsymbol{c}} \neq \varnothing$. 于是 $\boldsymbol{c} = \prec c^1, \cdots, c^n \succ$, 其中对所有 $i, c^i > 0$. 对每个 r, 有 n 个整数 $N^1, \cdots, N^n$ 使得 (图 8.6)

$$N^i/2^r \leqslant c^i \leqslant (N^i + 1)/2^r . \tag{4.3}$$

在下面, 设 $\boldsymbol{k} = \prec k^1, \cdots, k^n \succ, \boldsymbol{l} = \prec l^1, \cdots, l^n \succ$ 等等, 表示具有整数坐标的向量 (即 k^i 都为整数). 如果 $k_i < l_i$ 都所有 i 成立, 我们记其为 $\boldsymbol{k} < \boldsymbol{l}$. 如果 $\boldsymbol{N} = \prec N_1, \cdots, N_n \succ$, 则由 (4.3) 式和定义得到: 一旦 $\boldsymbol{k} < \boldsymbol{N}$ 则

$$\square_{(1/2^r)\boldsymbol{k}}^{(1/2^r)\boldsymbol{k}+\boldsymbol{1}/\boldsymbol{2}^r} \subset \square_{\boldsymbol{0}}^{\boldsymbol{c}} .$$

对任意 $\boldsymbol{k}$ 和 $\boldsymbol{l}$, 如果 $\boldsymbol{k} \neq \boldsymbol{1}$

$$\square_{(1/2^r)\boldsymbol{k}}^{(1/2^r)\boldsymbol{k}+\boldsymbol{1}/\boldsymbol{2}^r} \cap \square_{(1/2^r)\boldsymbol{l}}^{(1/2^r)\boldsymbol{l}+\boldsymbol{1}/\boldsymbol{2}^r} = \varnothing .$$

由于由 (4.2) 式 (及公理 $\mu2$)

$$\mu(\square_{(1/2^r)\boldsymbol{k}}^{(1/2^r)\boldsymbol{k}+\boldsymbol{1}/\boldsymbol{2}^r}) = \frac{1}{2^{nr}}$$

和

$$\bigcup_{\boldsymbol{k}<\boldsymbol{N}} \square_{(1/2^r)\boldsymbol{k}}^{(1/2^r)\boldsymbol{k}+\boldsymbol{1}/\boldsymbol{2}^r} \subset \square_{\boldsymbol{0}}^{\boldsymbol{c}}$$

我们得到

$$\mu(\square_{\boldsymbol{0}}^{\boldsymbol{c}}) \geqslant \frac{1}{2^{nr}} \times (\text{满足}0 \leqslant \boldsymbol{k} < \boldsymbol{N}\text{的 } \boldsymbol{k} \text{ 的个数}) .$$

容易看出有 $N_1 \cdot N_2 \cdot \cdots \cdot N_n$ 个这样的 $\boldsymbol{k}$, 故

$$\mu(\square_{\boldsymbol{0}}^{\boldsymbol{c}}) \geqslant \frac{1}{2^{nr}} \times (N_1 \cdots N_n) = \left(\frac{N_1}{2^r}\right) \cdots \left(\frac{N_n}{2^r}\right) .$$

根据 (4.3) 式, $N_i/2^r > c_i - 1/2^r$, 故而

$$\mu(\square_{\boldsymbol{0}}^{\boldsymbol{c}}) \geqslant \left(c^1 - \frac{1}{2^r}\right) \cdots \left(c^n - \frac{1}{2^r}\right) . \tag{4.4}$$

类似地,

$$\square_{\boldsymbol{0}}^{\boldsymbol{c}} \subset \bigcup_{\boldsymbol{k}<\boldsymbol{N}+2} \square_{(1/2^r)\boldsymbol{k}}^{(1/2^r)\boldsymbol{k}+\boldsymbol{1}/\boldsymbol{2}^r} ,$$

从而得到

$$\mu(\square_{\boldsymbol{0}}^{\boldsymbol{c}}) \leqslant \left(c^1 + \frac{2}{2^r}\right) \cdots \left(c^n + \frac{2}{c^r}\right) . \tag{4.5}$$

让在 (4.4) 式和 (4.5) 式中的 $r \to \infty$ 便证明了 (4.1) 式 .

在推导 (4.1) 式中我们利用了公理 $\mu4$. 查验我们的论证可以证明如果 μ' 满足 $\mu2$ 和 $\mu3$ 但不是$\mu4$, 我们则可以按同一方式进行论证, 然而有一点不同, 即我们不得不用固定的常数 $\mu'(\square_{\boldsymbol{0}}^{\boldsymbol{1}})$ 去乘每一个东西. 总合起来我们已证明了

命题 4.1 如果 μ 满足公理 $\mu1$ 到 $\mu4$, 则任意矩形的 μ 的值被惟一决定并由 (4.1) 式给出. 如果 μ' 满足 $\mu1$ 到 $\mu3$, 则对任意矩形 $\Box_{\boldsymbol{a}}^{\boldsymbol{b}}$, 有

$$\mu^1(\Box_{\boldsymbol{a}}^{\boldsymbol{b}}) = K\mu(\Box_{\boldsymbol{a}}^{\boldsymbol{b}}),\ 其中 K = \mu'(\Box_{\boldsymbol{0}}^{\boldsymbol{1}}) .$$

8.5 极小理论 (续)

我们现在要证明公式 (4.1) 可扩张到一个定义在 $\mathfrak{D}_{\min}$ 上满足公理 $\mu1$ 到 $\mu4$ 的惟一的 μ. 我们必须从实质上建立两个事实.

(1) 矩形的每个并集可以写成矩形的不交并.

这个事实让我们对每个 $A \in \mathfrak{D}_{\min}$ 按公理 $\mu1$ 决定 $\mu(A)$; 如果 A 是 $\Box_{\boldsymbol{a}_i}^{\boldsymbol{b}_i}$ 的不交并, 则令

$$\mu(A) = \sum \mu(\Box_{\boldsymbol{a}_i}^{\boldsymbol{b}_i}) .$$

由于 A 可以以另外的方式写成矩形的不交并, 除非我们建立了下面的事实, 否则这个公式并没有确切的定义.

(2) 如果 $A = \cup\Box_{\boldsymbol{a}_i}^{\boldsymbol{b}_i} = \cup\Box_{\boldsymbol{c}_i}^{\boldsymbol{d}_i}$ 为 A 的两个矩形不交并的表示, 则

$$\sum \mu(\Box_{\boldsymbol{a}_i}^{\boldsymbol{b}_i}) = \sum \mu(\Box_{\boldsymbol{c}_i}^{\boldsymbol{d}_i}) .$$

先引进一些记号.

定义 5.1 $\mathbb{E}^n$ 的一个铺砌 $\mathfrak{p}$ 是个互不相交的矩形的有限集合. 这个铺砌的地面是所有属于 $\mathfrak{p}$ 的矩形的并集, 记为 $|\mathfrak{p}|$.

如果 $\mathfrak{p} = \{\Box_{a_i}^{b_i}\}$, T 为平移, 我们令 $T\mathfrak{p} = \{T\Box_{a_i}^{b_i}\}$.

如果 $\mathfrak{p}$ 和 $\mathfrak{z}$ 是两个铺砌, 称 $\mathfrak{z}$ 细于 $\mathfrak{p}$ (记为 $\mathfrak{p} \prec \mathfrak{z}$) 是说如果 $\mathfrak{p}$ 中的每个矩形是 $\mathfrak{z}$ 中矩形的一个并集. 显然, 如果 $\mathfrak{p} \prec \mathfrak{v}, \mathfrak{z} \prec \mathfrak{p}$ 则 $\mathfrak{z} \prec \mathfrak{v}$. 还要注意 $\mathfrak{z} \prec \mathfrak{p}$ 蕴涵了 $|\mathfrak{p}| \subset |\mathfrak{z}|$

命题 5.1 设 $\mathfrak{p}$ 和 $\mathfrak{z}$ 为两个铺砌. 则存在第三个铺砌 $\mathfrak{v}$ 使得 $\mathfrak{v} \prec \mathfrak{p}$ 和 $\mathfrak{v} \prec \mathfrak{z}$.

证明 证明的想法极简单. $\mathfrak{p}$ 中或 $\mathfrak{z}$ 中每个矩形决定了 2^n 个超平面 (每个超平面包含矩形的一个面). 如果我们汇拢所有这些超平面, 它们将 "包" 住了许多个矩形. 我们设 $\mathfrak{v}$ 为在此汇拢的集合中那些不再包含更小的矩形的矩形组成. 图 8.7 展示了这种情形 ($n = 2$), 其中 $\mathfrak{p}$ 和 $\mathfrak{z}$ 每个只包含了一个矩形, 而 $\mathfrak{v}$ 包含了 9 个矩形.

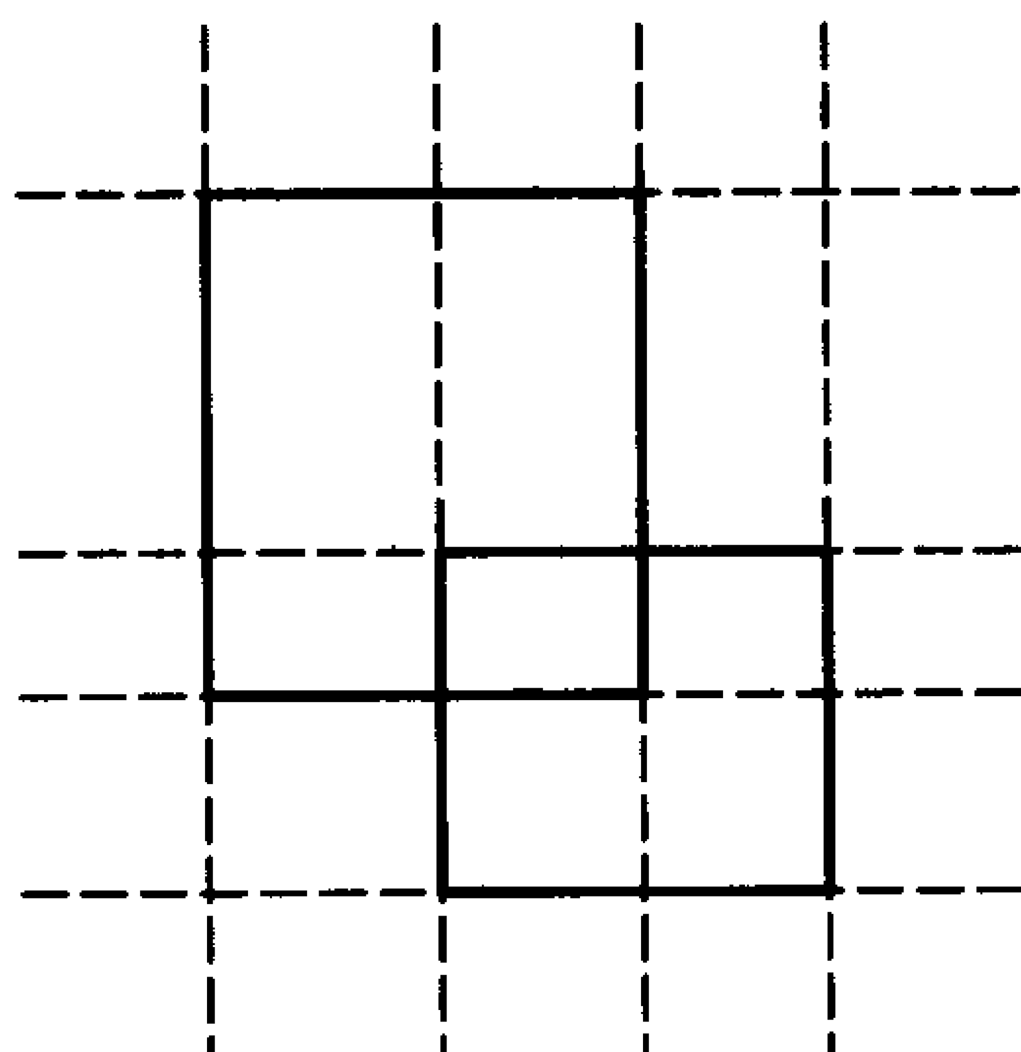


图 8.7

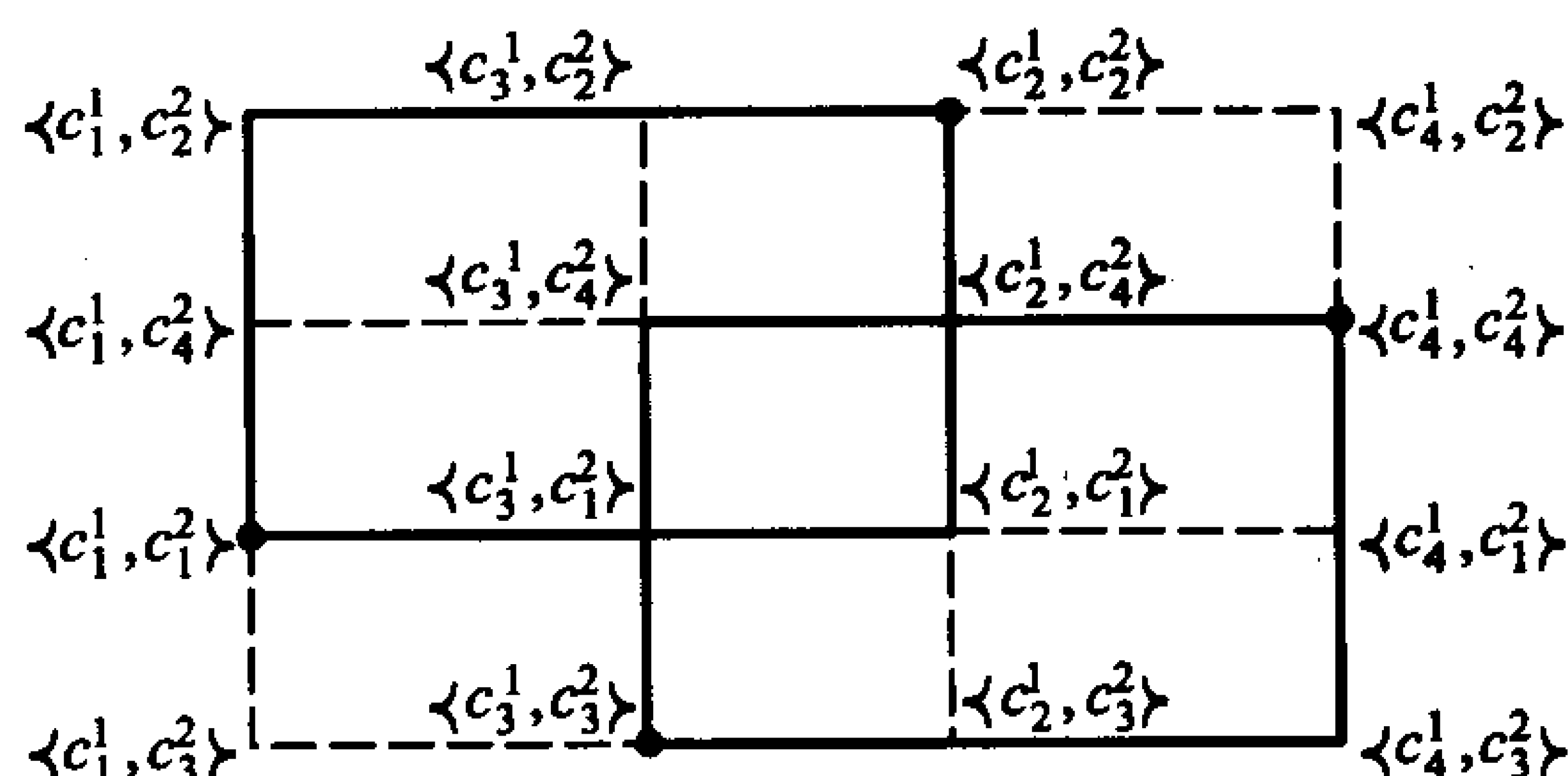


图 8.8

我们现在充满整个论证的细节. 设 $\boldsymbol{c}_1 =\prec c_1^1,\cdots,c_1^n \succ,\cdots,\boldsymbol{c}_k =\prec c_k^1,\cdots,c_k^n \succ$ 为所有那些出现在描述 $\mathfrak{p}$ 和 $\mathfrak{z}$ 中矩形的向量. (换句话说, 如果 $\Box_{\boldsymbol{a}}^{\boldsymbol{b}} \in \mathfrak{p}$ 或 $\in \mathfrak{z}$, 则 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 均在这些 $\boldsymbol{c}$ 中.) 设 $\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{k^n}$ 为形如 $\prec c_{i_1}^1,\cdots,c_{i_n}^n \succ$ 的向量, 其中 i_j 的值域各自独立地在 1 和 k 之间 (从而它们有 k^n 个). (见图 8.8 对 $n=2$ 和 $\mathfrak{p}$ 与 $\mathfrak{z}$ 各含一个矩形的情形.) 对每个 $\boldsymbol{d}_i$, 最多有一个最小的 $\boldsymbol{d}_{j(i)}$ 使 $\boldsymbol{d}_i < \boldsymbol{d}_{j(i)}$. 事实上, 如果

$$\boldsymbol{d}_i =\prec c_{i_1}^1,\cdots,c_{i_n}^n \succ ,$$

则令 $\boldsymbol{d}_{j(i)} =\prec c_{j_1}^1,\cdots,c_{j_n}^n \succ$, 其中

$$c_{j_l}^l = \min_{c_m^l > c_{i_l}^l} \{c_m^l\} .$$

设 $\mathfrak{v} = \{\Box_{\boldsymbol{d}_i}^{\boldsymbol{d}_{j(i)}}\}$. 于是 $\mathfrak{v}$ 比 $\mathfrak{p}$ 和 $\mathfrak{z}$ 细. 事实上, 如果 $\Box_{\boldsymbol{a}}^{\boldsymbol{b}} \in \mathfrak{p}$, 则 $\Box_{\boldsymbol{a}}^{\boldsymbol{b}} = \Box_{\boldsymbol{d}_\beta}^{\boldsymbol{d}_\alpha}$ 对某些 α 和 β, 且

$$\Box_{\boldsymbol{d}_\alpha}^{\boldsymbol{d}_\beta} = \bigcup_{\substack{\boldsymbol{d}_\alpha \leqslant \boldsymbol{d}_i \\ \boldsymbol{d}_{j(i)} \leqslant \boldsymbol{d}_\beta}} \Box_{\boldsymbol{d}_i}^{\boldsymbol{d}_{j(i)}} . \tag{5.1}$$

要明白这点, 只要看到如果 $\boldsymbol{x} \in \Box_{\boldsymbol{d}_\alpha}^{\boldsymbol{d}_\beta}$ 则 $\boldsymbol{d}_\alpha \leqslant \boldsymbol{x} < \boldsymbol{d}_\beta$. 选取一个最大的 $\boldsymbol{d}_i \leqslant \boldsymbol{x}$. 于是 $\boldsymbol{d}_i \leqslant \boldsymbol{x} < \boldsymbol{d}_{j(i)}$, 故 $\boldsymbol{x} \in \Box_{\boldsymbol{d}_i}^{\boldsymbol{d}_{j(i)}}$. 这便证明了此命题. 我们在后面要用到我们所构造的这个 $\mathfrak{v}$ 的特殊形式, 以寻找更多的信息. □

现在能够证明 (1) 和 (2) 了.

引理 5.1 *设 $\mathfrak{p}_1,\cdots,\mathfrak{p}_l$ 为铺砌. 于是存在一个铺砌子使得 $|\mathfrak{z}| = |\mathfrak{p}_1|\cup\cdots\cup|\mathfrak{p}_l|$.*

证明 反复应用命题 5.1, 我们便能选取一个铺砌 $\mathfrak{v}$, 细于所有的 $\mathfrak{p}_i$. 于是每个 $|\mathfrak{p}_i|$ 是 $\mathfrak{v}$ 中某些适当的矩形的并. 令 $\mathfrak{z}$ 为所有那些出现在所有 $\mathfrak{p}_i$ 中的矩形的集合. 于是 $|\mathfrak{z}| = |\mathfrak{p}_1| \cup \cdots \cup |\mathfrak{p}_k|$. □

特别地, 我们已经证好了 (1). 更一般, 我们也已证明了每个 $A \in \mathfrak{D}_{\min}$ 均具有形式 $A = |\mathfrak{p}|$, 其中 $\mathfrak{p}$ 为某个适当的铺砌. 我们现在要转而讨论 (2).

引理 5.2 设 $c_1^1 < \cdots < c_{r_1}^1, c_1^2 < \cdots < c_{r_2}^2, \cdots, c_1^n < \cdots < c_{r_n}^n$ 为 n 个数的序列. 于是

$$\mu\Box_{\prec c_1^1,\cdots,c_1^n\succ}^{\prec c_{r_1}^1,\cdots,c_{r_n}^n\succ} = \sum_{\substack{1\leqslant i_1<r_1\\ \vdots \\ 1\leqslant i_n<r_n}} \mu\Box_{\prec c_{i_1}^1,\cdots,c_{i_n}^n\succ}^{\prec c_{i_1+1}^1,\cdots,c_{i_n+1}^n\succ}.$$

证明 事实上, $c_{r_i}^i - c_1^i = c_2^i - c_1^i + c_3^i - c_2^i + \cdots + c_{r_i}^i - c_{r_{i-1}}^i$, 故当我们乘出所有因子时引理便由 (4.1) 式得出. □

现在证明 (2). 设 $\mathfrak{p} = \{\Box_{a_i}^{b_i}\}, \mathfrak{z} = \{\Box_{e_k}^{f_k}\}$, 其中 $A = |\mathfrak{p}| = |\mathfrak{z}|$. 设 $\mathfrak{v}$ 为我们在命题 5.1 的证明中所构造的铺砌. 设 $\mathfrak{s} = \{\Box_{d_l}^{d_m}\}$ 为 $\mathfrak{v}$ 中使 $\Box_{d_\alpha}^{d_\beta} \subset |\mathfrak{p}| = |\mathfrak{z}|$ 的那些矩形 $\Box_{d_\alpha}^{d_\beta}$ 的汇集. 那么要证明 (2) 只要证明

$$\sum \mu(\Box_{d_l}^{d_m}) = \sum \mu(\Box_{a_i}^{b_i}) = \sum \mu(\Box_{e_k}^{f_k}) \tag{5.2}$$

就可以了. 现在每个矩形 $\Box_{a_i}^{b_i}$ 按 (5.1) 式被分解为矩形 $\Box_{d_l}^{d_{j(l)}}$, 即 $a_i = d_\alpha, b_i = d_\beta$, 等等.

由这些 d 的构造知, 这完全是引理 5.2 中所描述的那种分解类型. 因此 (5.1) 式蕴函了

$$\mu(\Box_{d_\alpha}^{d_\beta}) = \sum_{\substack{d_\alpha \leqslant d_i\\ d_{j(i)} < d_\beta}} \mu(\Box_{d_i}^{d_{j(i)}}).$$

对所有 $\Box_{a_i}^{b_i}$ 取和 (并对 $\Box_{c_j}^{d_j}$ 同样进行) 便证明了 (5.2) 式. 因此我们可叙述为:

定理 5.1 每个 $A \in \mathfrak{D}_{\min}$ 可写为 $A = |\mathfrak{p}|$. 数 $\mu(A) = \sum_{\Box\in\mathfrak{p}} \mu(\Box)$ 不依赖于 $\mathfrak{p}$ 的选取. 因此我们得到 $\mathfrak{D}_{\min}$ 上一个有明确定义的函数 μ. 它满足公理 $\mu 1$ 到 $\mu 4$. 如果 μ' 是 $\mathfrak{D}_{\min}$ 上另一个满足 $\mu 2$ 和 $\mu 3$ 的函数, 则 $\mu'(A) = K\mu(A)$, 其中 $K = \mu'(\Box_0^1)$.

证明 定理的最后两个结论的证明是容易的, 留给读者当作习题. □

8.6 可度集合

定理 5.1 表明我们的公理不是个空壳. 然而它并没有提供一个满意的理论, 因为 $\mathfrak{D}_{\min}$ 所含有的集合太少了. 特别, 它不满足要求 (iii), 这是因为除了一些极为特殊的情形外, $\mathfrak{D}_{\min}$ 在旋转下不是不变的. 我们现在要修补它, 办法是重复使用第 4 节的论证; 我们要用那些已知其 μ 的集合, 即包含在 $\mathfrak{D}_{\min}$ 中的集合去力图逼近更为一般的集合. 这种思想可追溯到阿基米德, 他使用这种思想去求平面图形的面积 .

定义 6.1 设 A 为 $\mathbb{E}^n$ 的任意子集. 如果 $|\mathfrak{p}| \subset A$ 则称 $\mathfrak{p}$ 为 A 的一个内铺砌.如果 $A \subset |\mathfrak{z}|$ 则称 $\mathfrak{z}$ 为 A 的一个外铺砌 .

我们列出数个显然的事实.

$$\text{如果 } |\mathfrak{p}| \subset A \subset |\mathfrak{z}|, \text{ 则 } \mu(\mathfrak{p}) \leqslant \mu(\mathfrak{z}) . \tag{6.1}$$

$$\text{如果 } |\mathfrak{p}| \subset A \subset |\mathfrak{z}|, \text{ 则 } |T\mathfrak{p}| \subset TA \subset |T\mathfrak{z}| . \tag{6.2}$$

$$\text{如果 } A_1 \cap A_2 = \varnothing \text{且 } |\mathfrak{p}_1| \subset A_1, |\mathfrak{p}_2| \subset A_2 ,$$

$$\text{则 } \mathfrak{p}_1 \cup \mathfrak{p}_2 \text{ 是 } A_1 \cup A_2 \text{ 的内铺砌} . \tag{6.3}$$

定义 6.2 对 $\mathbb{E}^n$ 中任一有界子集 A, 设

$$\mu^*(A) = \operatorname*{lub}_{|\mathfrak{p}| \subset A} \mu(|\mathfrak{p}|) ,$$

称其为 A 的内容度 .又设

$$\overline{\mu}(A) = \operatorname*{glb}_{A \subset |\mathfrak{z}|} \mu(|\mathfrak{z}|) ,$$

称其为 A 的外容度 .

注意, 由于 A 是有界的, 故存在一个 $\mathfrak{z}$ 满足 $A \subset |\mathfrak{z}|$. 这表明 $\overline{\mu}(A)$ 有定义. 这个连同 (6.1) 式表明 $\mu^*(A)$ 有定义并且

$$\mu^*(A) \leqslant \overline{\mu}(A) . \tag{6.4}$$

定义 6.3 称集合 A 是可度的是说如果有 $\mu^*(A) = \overline{\mu}(A)$. 我们称 $\mu^*(A) = \overline{\mu}(A)$ 为 A 的容度, 并记为 $\mu(A)$.

每个 $A \in \mathfrak{D}_{\min}$ 是可度的. 事实上, 如果 $A = |\mathfrak{v}|$, 则 $\mathfrak{v}$ 同时是 A 的内铺砌和外铺砌. 因此 $\mu^*(A) = \overline{\mu}(A) = \mu(|\mathfrak{v}|)$, 从而 $\mu(A)$ 的新定义与老的一致.

我们的下一个直接目标是证明所有的可度集合的汇集满足公理 $\mathfrak{D}1$ 到 $\mathfrak{D}3$.

命题 6.1 集合 A 可度的充要条件是它的边界是可度的并且有容度 0.

证明 假定 A 是可度的. 对任何 $\delta > 0$, 可以找到一个内铺砌 $\mathfrak{p}$ 和一个外铺砌 $\mathfrak{z}$ 使得 $\mu(\mathfrak{z}) - \mu(\mathfrak{p}) < \delta/2$. 我们要用一个满足 $|\mathfrak{p}'| \subset \operatorname{int} A$ 的比 $\mathfrak{p}$ 更靠近的铺砌去替换 $\mathfrak{p}$. 要做到这点, 我们选取一个小的数 η 并以 $\square_{a+\eta(b-a)}^{b-\eta(b-a)}$ 去替代 $\mathfrak{p}$ 中的每个矩形 $\square_a^b$. 以 $\mathfrak{p}_\eta$ 记所有这些矩形的集合. 于是 $|\mathfrak{p}_\eta| \subset \operatorname{int}|\mathfrak{p}|$, 从而 $|\mathfrak{p}_\eta| \subset \operatorname{int} A$. 进一步还有 $\mu(|\mathfrak{p}_\eta|) = (1-2\eta)^n\mu(|\mathfrak{p}|)$, 这是因为 $(1-2\eta)$ 是 $\mathfrak{p}$ 中每个矩形的每边的减少因子. 相似地, 用稍大一点的满足 $A \subset \operatorname{int}|\mathfrak{z}_\eta|$ 的 $\mathfrak{z}_\eta$ 替换 $\mathfrak{z}$, 并有 $\mu(|\mathfrak{z}_\eta|) \leqslant (1+2\eta)^n\mu(|\mathfrak{z}|)$. 选取 η 充分小, 因而可使得 $\mu(|\mathfrak{z}_\eta|) - \mu(|\mathfrak{p}_\eta|) < \delta$. 设 $\mathfrak{v}$ 是比 $\mathfrak{z}_\eta$ 和 $\mathfrak{p}_\eta$ 还细的一个铺砌, 满足 $|\mathfrak{v}| = |\mathfrak{z}_\eta|$. 设 $\mathfrak{s} \subset \mathfrak{v}$ 由 $\mathfrak{v}$ 中那些位于 $\operatorname{int} A$ 中的矩形组成. 于是 $|\mathfrak{v}| = |\mathfrak{z}_\eta| \supset |\mathfrak{s}| \supset |\mathfrak{p}_\eta|$, 故 $\mu(|\mathfrak{v}|) - \mu(|\mathfrak{s}|) \leqslant \delta$. 但 $\partial A \subset |\mathfrak{v} - \mathfrak{s}|$, 从而 $\overline{\mu}(\partial A) \leqslant \mu(|\mathfrak{v} - \mathfrak{s}|) = \mu(|\mathfrak{v}|) - \mu(|\mathfrak{s}|) < \delta$. 换句话说, $\overline{\mu}(\partial A) = 0$.

反之, 假定 ∂A 的容度为 0. 设 $\mathfrak{s}$ 为 ∂A 的一个外铺砌使 $\mu(|\mathfrak{s}|) < \varepsilon$. 设 $\mathfrak{v}$ 为细于 $\mathfrak{s}$ 的一个铺砌使得 $A \subset |\mathfrak{v}|$. 设 $\mathfrak{p} \subset \mathfrak{v}$ 由那些包含于 A 内的那些矩形组成. 让 $\mathfrak{z} \subset \mathfrak{v}$ 由那些在 $|\mathfrak{p}| \cup |\mathfrak{s}|$ 中的矩形组成. 于是 $\mu(|\mathfrak{z}|) \leqslant \mu(|\mathfrak{p}|) + \mu(|\mathfrak{s}|) < \mu(|\mathfrak{p}|) + \varepsilon$. 进一步有 $A \subset |\mathfrak{z}|$. 事实上, 让 $x \in A$, 则对某个 $\square \in \mathfrak{v}$ 有 $x \in \square$. 如果 $\square \cap \partial A \neq \varnothing$, 那么 $\square \cap |\mathfrak{s}| \neq \varnothing$, 故由于 $\mathfrak{v}$ 是 $\mathfrak{s}$ 的一个细分便有 $\square \subset |\mathfrak{s}|$. 如果 $\square \cap \partial A = \varnothing$, 则 $\square$ 中每个点都在 A 中, 故 $\square \subset |\mathfrak{p}|$. 这样, 我们已经构造了 $\mathfrak{p}$ 和 $\mathfrak{z}$, 它们满足 $|\mathfrak{p}| \subset A \subset |\mathfrak{z}|$ 和 $\mu(\mathfrak{z}) - \mu(\mathfrak{p}) < \varepsilon$. 因为我们对任意 ε 都可这样做, 故 A 是可度的. □

命题 6.2 任何有限个具零容度集合的并仍具零容度. 如果 $A \subset B$, 而 B 具零容度, 则 A 亦然,

证明 证明是显然的. □

定理 6.1 设 $\mathfrak{D}_{\text{con}}$ 表示所有可度集合的汇集. 于是 $\mathfrak{D}_{\text{con}}$ 满足公理 $\mathfrak{D}1$ 到 $\mathfrak{D}3$, 并且由定义 6.3 给出的 μ 满足 $\mu1$ 到 $\mu4$. 如果 μ' 为 $\mathfrak{D}_{\text{con}}$ 上满足 $\mu1$ 到 $\mu3$ 的任何其他一个函数, 则 $\mu' = K\mu$, 其中 $K = \mu'(\square_0^1)$.

证明 让我们来验证这些公理.

$\mathfrak{D}1$. 对任意 A 和 B，

$$\partial(A\cup B)\subset\partial A\cup\partial B \quad \text{和} \quad \partial(A\cap B)\subset\partial A\cup\partial B\,.$$

由命题 6.1, 如果 A 和 B 均可度, 则 ∂A 和 ∂B 具容度零. 因此由命题 6.2, $\partial A\cup\partial B, \partial(A\cup B)$ 和 $\partial(A\cap B)$ 都具零容度, 从而 $A\cup B$ 和 $A\cap B$ 可度 .

$\mathfrak{D}2$. 立即由 (6.1) 式推导出 .

$\mathfrak{D}3$. 显然 .

$\mu2$. 如果 A_1 和 A_2 是可度的, 我们可以找到内铺砌 $\mathfrak{p}_1$ 和 $\mathfrak{p}_2$ 使得 $\mu(A_1)-\mu(|\mathfrak{p}_1|)<\varepsilon/2$ 和 $\mu(A_2)-\mu(|\mathfrak{p}_2|)<\varepsilon/2$. 如果 $A_1\cap A_2=\varnothing$, 则 $\mathfrak{p}_1\cup\mathfrak{p}_2$ 是 $A_1\cup A_2$ 的内铺砌, 故

$$\mu(A_1\cup A_2)\geqslant\mu(A_1)+\mu(A_2)\,.$$

另一方面, 设 $\mathfrak{z}_1$ 和 $\mathfrak{z}_2$ 分别为 A_1 和 A_2 的外铺砌, 并满足 $\mu(|\mathfrak{z}_1|)<\mu(A_1)+\varepsilon/2$ 和 $\mu(|\mathfrak{z}_2|)<\mu(A_2)+\varepsilon/2$.

设 $\mathfrak{v}$ 是一个满足 $|\mathfrak{v}|=|\mathfrak{z}_1|\cup|\mathfrak{z}_2|$ 的铺砌. 那么 $\mathfrak{v}$ 是 $A_1\cup A_2$ 的一个外铺砌, 且满足 $\mu(|\mathfrak{v}|)\leqslant\mu(|\mathfrak{z}_1|)+\mu(|\mathfrak{z}_2|)$. 因此 $\mu(A_1\cup A_2)\leqslant\mu(|\mathfrak{v}|)\leqslant\mu(A_1)+\mu(A_2)+\varepsilon$, 或者 $\mu(A_1\cup A_2)\leqslant\mu(A_1)+\mu(A_2)$. 这两个不等式合起来给出了 $\mu2$.

$\mu1$. 显然 .

$\mu3$. 由 (6.2) 式和定义 6.3 得到 .

$\mu4$. 已知 .

定理的第二部分可由定理 5.1 和定义 6.3 推出. 事实上, 我们知道 $\mu'(|\mathfrak{p}|)=K\mu(|\mathfrak{p}|)$, (6.1), 连同公理 $\mu2$ 意味着 $\mu'(|\mathfrak{p}|)\leqslant\mu'(A)\leqslant\mu'(|\mathfrak{z}|)$. 由于我们可选 $\mathfrak{p}$ 和 $\mathfrak{z}$ 任意靠近 A(在 μ 下), 定理便得证 . □

注 注意到我们实际上所证明的要比定理 6.1 所叙述的要多一点, 它对我们来说是有用的. 就是说, 我们已经证明了, 如果 $\mathfrak{D}$ 为任何满足 $\mathfrak{D}1$ 到 $\mathfrak{D}3$ 的集合的汇集, 使得 $\mathfrak{D}_{\min}\subset\mathfrak{D}\subset\mathfrak{D}_{\text{con}}$，并且如果 $\mu':\mathfrak{D}\to\mathbb{R}$ 满足 $\mu1$ 到 $\mu3$, 则对所有 $A\in\mathfrak{D}$ 有 $\mu'(A)=K\mu(A)$, 其中 $K=\mu'(\Box_0^1)$.

8.7 何时可度?

我们要建立一些准则以决定一个已知集合是否可度. 回忆一个中心在 $\boldsymbol{x}$ 半径为 r 的闭球是由

$$B_{\boldsymbol{x}}^r=\{\boldsymbol{y}:\|\boldsymbol{y}-\boldsymbol{x}\|\leqslant r\} \tag{7.1}$$

给出的. 我们看到有

$$B_x^r \subset \square_{x-r1}^{x+(r+\varepsilon)1}, \text{ 其中 } \varepsilon \text{为任意的} \varepsilon > 0\,, \tag{7.2}$$

和

$$\square_{x-r1}^{x+r1} \subset B_x^{\sqrt{n}r}\ . \tag{7.3}$$

(见图 8.9)

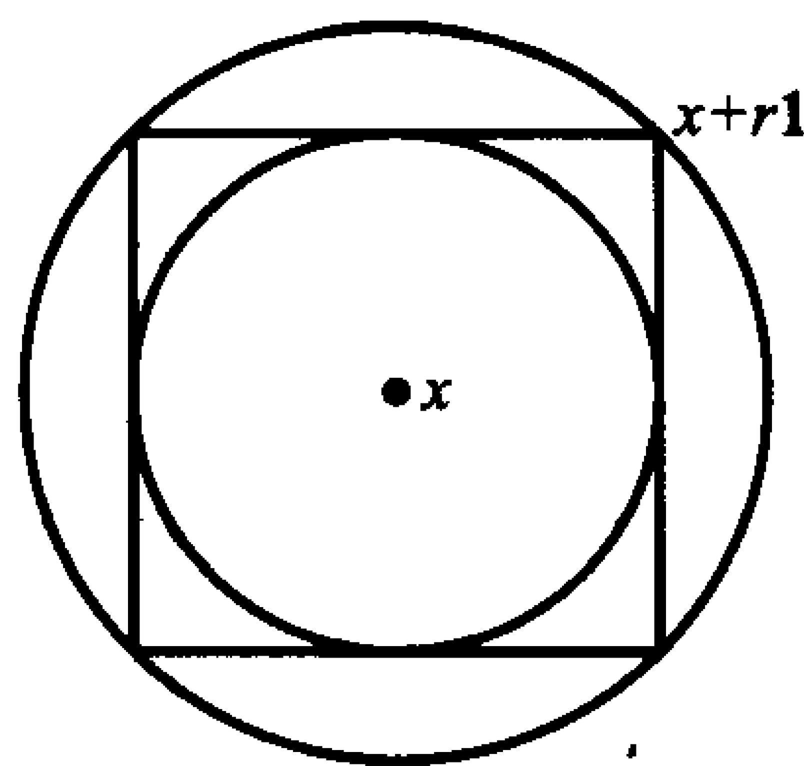


图 8.9

如果联合 (7.2) 和 (7.3) 式, 我们看到任何在球 B 中的立方体C 有 $\overline{\mu}(B) \leqslant 2^n(\sqrt{n})^n\mu(C)$, 而任何在立方体 C 内的球 B 有 $\mu(c) < 3^n(\sqrt{n})^n\overline{\mu}(B)$.

引理 7.1 *设 A 为 $\mathbb{E}^n$ 的一个子集. 则 A 有零容度当且仅当对任意 $\varepsilon > 0$, 存在有限个覆盖 A 的球 $\{B_i\}$ 使得 $\sum \overline{\mu}(B_i) < \varepsilon$.*

证明 如果我们有这样一组覆盖的球, 则由上面的注知道, 我们可扩大每个球为一个矩形从而得到一个铺砌 $\mathfrak{p}$, 使得 $A \subset |\mathfrak{p}|$ 及 $\mu(|\mathfrak{p}|) < 3^n(\sqrt{n})^n\varepsilon$. 因此, 如果我们总能找到这些球 $\{B_i\}$ 则 $\overline{\mu}(A) = 0$.

反之, 假设 A 有容度零. 于是对任意 δ, 我们可以找到一个外铺砌 $\mathfrak{p}$ 满足 $\mu(|\mathfrak{p}|) < \delta$. 对此铺砌中每一个矩形 $\square$, 用第 4 节的论证, 我们可以找到有限个立方体覆盖 $\square$, 而且它们的总容度可以按我们的要求任意靠近 $\mu(\square)$, 譬如 $< 2\mu(\square)$. 对每个 $\square \in \mathfrak{p}$ 都这样做我们便有有限个立方体 $\{\square_i\}$, 它们覆盖了 A 并有小于 2δ 的总容度. 那么由此引理前面的评注知道, 每个立方体 $\square_i$ 位于一个球 B_i 内, 使得 $\mu(\square_i) < 2^n(\sqrt{n})^n\overline{\mu}(B_i)$, 从而我们有了一个 A 的用球 B_i 的覆盖, 使得 $\sum \overline{\mu}(B_i) < 2^{n+1}(\sqrt{n})^n\delta$. 如果取 $\delta = \varepsilon/2^{n+1}(\sqrt{n})^n$, 我们便有了想要的球的集合, 从而证明了引理 . □

回想起称一个从 $U \subset \mathbb{E}^n \to \mathbb{E}^n$ 的映射 φ 满足利普希茨条件是说, 如果存在常数 K(称作利普希茨常数) 使得

$$\|\varphi(y) - \varphi(x)\| < K\|y - x\|\ . \tag{7.4}$$

命题 7.1 *设 A 为容度零的集合, 并满足 $\overline{A} \subset U$, 又设 $\varphi : U \to \mathbb{E}^n$ 满足利普希茨条件. 则 $\varphi(A)$ 具容度零 .*

证明 证明由对两部分都应用引理 7.1 构成. 因为 A 具零容度, 故对 $\varepsilon > 0$ 可以找到有限个球覆盖了 A, 而其外容度小于 ε/K^n. 由 (7.4) 式, $\varphi(B_x^r) \subset B_{\varphi(x)}^{kr}$, 故覆盖 A 的球的像覆盖了 $\varphi(A)$, 从而总体积小于 ε . □

回忆下面的事实: 如果 φ 是由开集 U 到 $\mathbb{E}^n$ 的 (连续) 可微映射, 则在 U 的任一紧子集上 φ 满足利普希茨条件 .

作为命题 7.1 的推论, 我们可以叙述下面的命题:

命题 7.2 设 φ 是定义于开集 U 上的连续可微映射, 并设 A 为一容度零的有界集, 满足 $\overline{A} \subset U$. 则 $\varphi(A)$ 具有零容度 .

设 A 为 $\mathbb{E}^n$ 中任一完全位于由 $x^n = 0$ 给出的子空间内的紧子集, 则 A 具有零容度. 事实上, 对某个充分大的固定的 r, 集合 A 包含在矩形

$$\square_{\prec -r,\cdots,-r,0\succ}^{\prec r,\cdots,r,\varepsilon\succ}$$

中, 其中 $\varepsilon > 0$ 任意数, 此矩形具有任意小的体积 .

现设 ψ 为 $V \subset \mathbb{E}^{n-1}$ 到 $\mathbb{E}^n$ 的连续可微映射, 它由

$$\prec y^1,\cdots,y^{n-1} \succ \mapsto \prec \psi^1(y^1,\cdots,y^{n-1}),\cdots,\psi^n(y^1,\cdots,y^{n-1}) \succ$$

给出. 设 B 为 $\mathbb{E}^{n-1}$ 中任意有界子集, 满足 $\overline{B} \subset V$. 那么, 我们可写成 $\psi(B) = \varphi(A)$, 其中 A 为 $\mathbb{E}^n$ 中形如 $(y,0), y \in B$ 的点的集合, 而 φ 是个可微映射, 定义为

$$\varphi(x^1,\cdots,x^n) = \prec \psi^1(x^1,\cdots,x^{n-1}),\cdots,\psi^n(x^1,\cdots,x^{n-1}) \succ .$$

由命题 7.2 知, $\mu(\psi(B)) = 0$. 因此有

命题 7.3 设 ψ 为 $V \subset \mathbb{E}^{n-1}$ 到 $\mathbb{E}^n$ 中的可微映射, 并设 B 为满足 $\overline{B} \subset V$ 的有界集. 于是 $\psi(B)$ 的容度为零 .

我们因而恢复了第 1 节的要求 (v) .

命题 7.3 和 6.1 式的一个直接推论是 :

命题 7.4 设 $A \subset \mathbb{E}^n$ 使 $\partial A \subset \cup\psi_i B_i)$, 其中 ψ_i 和 B_i 如命题 7.3 中的 ψ 和 B. 则 A 是可度的 .

这表明每个 "我们可以画出来的" 集合都是可度的 .

习题

证明每个球是可度的 .

8.8 在线性畸变下的行为

我们继续推导出命题 7.1 的各种推论 .

命题 8.1 *设 φ 为 $U \to \mathbb{E}^n$ 的一个一对一的映射, 并且满足利普希茨条件, 以及 φ^{-1} 为连续. 如果 $\overline{A} \subset U$ 可度, 则 $\varphi(A)$ 也可度 .*

证明 因为 A 可度, 故 ∂A 为零容度. 由 φ 所满足的条件知 $\partial\varphi(A) = \varphi(\partial A)$. 因此 $\partial\varphi(A)$ 具有零容度, 从而 $\varphi(A)$ 可度 . □

命题 8.1 的直接推论是 :

命题 8.2 *如果 L 是 $\mathbb{E}^n$ 的线性变换. 只要 A 为可度则 LA 可度 .*

证明 如果 L 非异, 则直接用命题 8.1. 如果 L 奇异, 它将 $\mathbb{E}^n$ 整个映射成一个真子空间. 任何一个这样的子空间均含在 $\{\boldsymbol{x} : x^n = 0\}$ 适当线性变换下的像中, 从而 $\mu(LA) = 0$ 对任何可度的 A 成立 . □

定理 8.1 *设 L 为 $\mathbb{E}^n$ 的一个线性变换. 于是对任意的可度集 A 有*

$$\mu(LA) = |\det L|\mu(A) . \tag{8.1}$$

证明 我们只局限在非异的 L 情形, 这是因为我们已经对 $\det L = 0$ 的情形检验过 (8.1) 式. 如果 L 非异, 则 L 将可度集的类带到自己. 让我们对每个 $A \in \mathfrak{D}_{\text{con}}$ 定义 μ' 为 $\mu'(A) = \mu(LA)$. 我们断定 μ' 满足 $\mathfrak{D}_{\text{con}}$ 上的公理 $\mu 1$ 到 $\mu 3$.

事实上, $\mu 1$ 和 $\mu 2$ 的真实性是明显的, 而 $\mu 3$ 由下面的事实得到, 即对任意平移 T_v, 我们有 $T_{Lv}L = LT_v$, 故

$$\mu'(T_v A) = \mu(LT_i A) = \mu(T_{Lv}LA) = \mu(LA) = \mu'(A) .$$

由定理 5.2 我们有结论

$$\mu' = k_L \mu ,$$

其中 k_L 是依赖于 L 的某个常数. 我们必须证明 $k_L = |\det L|$.

我们首先看到, 如果 O 是个正交变换, 则

$$\mu(OA) = \mu(A) .$$

实际上, 我们已知 $\mu(OA) = k_O\mu(A)$. 如果取 A 为单位球 B_0^1, 则 $OB_0^1 = B_0^1$, 故 $k_0 = 1$.

其次我们看到 $\mu(L_1L_2A) = k_{L_1}\mu(L_2A) = k_{L_1}k_{L_2}\mu(A)$, 故

$$k_{L_1L_2} = k_{L_1}k_{L_2} .$$

现在我们想起任何一个非异的 L 可以写为 $L = PO$, 其中 P 为一个正定的自伴算子而 O 为正交. 因此 $k_L = k_P$ 且 $|\det L| = |\det P|\,|\det O| = |\det P|$, 故我们只需对正定的自伴线性变换验证 (8.1) 式, 任何这样的 P 可以写为 $P = O_1DO_1^{-1}$, 其中 O_1 是正交变换, D 是对角变换. 因为 P 正定, D 的所有特征值也是正定的. 由于 $\det P = \det D, k_P = k_D$, 故只需要对 L 为由对角矩阵给出的并具有正特征值 $\lambda_1, \cdots, \lambda_n$ 的线性变换 L 验证 (8.1) 式即可. 但这时有 $L\square_0^1 = \square_0^{\prec\lambda_1,\cdots,\lambda_n\succ}$, 故

$$\mu'(\square_0^1) = \mu(\square_0^{\prec\lambda_1,\cdots,\lambda_n\succ}) = \lambda_1\cdots\lambda_n = |\det L| ,$$

从而 (8.1) 式成立. □

习题

设 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 为 $\mathbb{E}^n$ 中的向量, 我们说 $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n$ 张成的平行六面体是指所有形如 $\sum_{i=1}^n x^i\boldsymbol{v}_i, 0 \leqslant x^i \leqslant 1$. 证明它的容度是 $|\det((\boldsymbol{v}_i, \boldsymbol{v}_j))|^{1/2}$.

8.9 积分的公理

迄今我们已证明有一个惟一的定义于 $\mathbb{E}^n$ 中集合的一个很大的汇集上的 μ. 然后除去一些极特殊的情形外, 我们并没有计算 μ 的一个有效方法. 要弥补这点缺憾, 我们必须引进一个积分理论. 首先介绍一些记号：

定义 9.1 设 f 为 $\mathbb{E}^n$ 上任一实函数. 记为 supp f 的 f 的支集是指使 f 不为零的点的集合的闭包; 即

$$\text{supp } f = \overline{\{\boldsymbol{x} : f(\boldsymbol{x}) \neq 0\}} .$$

我们注意到有

$$\text{supp } (f+g) \subset \text{supp } f \cup \text{supp } g \tag{9.1}$$

和

$$\text{supp } fg \subset \text{supp } f \cap \text{supp } g . \tag{9.2}$$

我们称 f 具有紧支集是说 supp f 为紧集. 公式 (9.1)[以及 g 为常数时的公式 (9.2)] 表明所有有紧支集的函数构成了一个向量空间 .

设 T 为 $\mathbb{E}^n$ 到自己的一个一对一变换. 对任意函数 f 我们以 Tf 表示函数

$$(Tf)(\boldsymbol{x}) = f(T^{-1}\boldsymbol{x}) . \tag{9.3}$$

我们看到, 如果 T 和 T^{-1} 为连续, 则

$$\text{supp } Tf = T\text{supp } f . \tag{9.4}$$

定义 9.2 设 A 为 $\mathbb{E}^n$ 的子集. 以 e_A 记 A 的特征函数, 即函数

$$e_A(\boldsymbol{x}) = \begin{cases} 1, & \text{如果 } \boldsymbol{x} \in A , \\ 0, & \text{如果 } \boldsymbol{x} \notin A . \end{cases} \tag{9.5}$$

我们有

$$e_{A_1 \cap A_2} = e_{A_1} \cdot e_{A_2} , \tag{9.6}$$

$$e_{A_1 \cup A_2} = e_{A_1} + e_{A_2} - e_{A_1 \cap A_2} , \tag{9.7}$$

$$\text{supp } e_A = \overline{A} , \tag{9.8}$$

并且对 $\mathbb{E}^n$ 到自己的一一映射 T 有

$$Te_A = e_{TA} . \tag{9.9}$$

我们所说的 $\mathbb{E}^n$ 上的积分论是指一个函数的汇集 $\mathfrak{F}$, 连同一个规则 $\int$, 它对每个 $f \in \mathfrak{F}$ 指定一个实数 $\int f$, 其遵从下列公理 :

$\mathfrak{F}1$. $\mathfrak{F}$ 是一个所有具紧支集的有界函数空间中的一个向量子空间 .

$\mathfrak{F}2$. 如果 $f \in \mathfrak{F}$,T 为平移, 则 $Tf \in \mathfrak{F}$.

$\mathfrak{F}3$. $e_\square$ 属于 $\mathfrak{F}$, 其中 $\square$ 为任意的矩形 .

$\int 1$. $\int$ 是 $\mathfrak{F}$ 上的线性函数 .

$\int 2$. $\int Tf = \int f$, 其中 T 为任意变换 .

$\int 3$. 如果 $f \geqslant 0$, 则 $\int f \geqslant 0$.

$\int 4$. $\int e_{\square_0^1} = 1$.

我们可看出, 这些公理蕴涵着 $\mathfrak{F}$ 包含了所有形如 $e_{\square_1} + e_{\square_2} + \cdots + e_{\square_k}$ 的函数, 其中 $\square_1, \cdots, \square_k$ 为任意的矩形. 特别, 对任意的铺砌 $\mathfrak{p}$, 函数 $e_{|\mathfrak{p}|}$ 必属于 $\mathfrak{F}$.

还有, 从 $\int 3$ 我们立刻有一个更强的形式 :

$\int 3'$. $f \leqslant g \Rightarrow \int f \leqslant \int g$, 这是由于有 $g - f \geqslant 0$.

命题 9.1 设 $\mathfrak{F}, \int$ 为一满足公理 $\mathfrak{F}1$ 到 $\mathfrak{F}3$ 和 $\int 1$ 到 $\int 4$ 的一个系统. 于是对任何满足 $e_A \in \mathfrak{F}$ 的可度集 A 有

$$\int e_A = \mu(A) \tag{9.10}$$

以及对每个 $f \in \mathfrak{F}$,

$\int 5$. $|\int f| \leqslant \|f\|_\infty \mu(\mathrm{supp}\, f)$.

证明 这些公理保证了对每个 $A \in \mathfrak{D}_{\min}$ 有 $e_A \in \mathfrak{F}$, 并且 $v(A) = \int e_A$ 满足 $\mu 1$ 到 $\mu 4$. 因此, $\int e_A = \mu(A)$ 由 μ 的惟一性 (命题 4.1) 对每个 $A \in \mathfrak{D}_{\min}$ 成立. 由此得到, 如果 A 是一个使 $e_A \in \mathfrak{F}$ 的可度集, 并且如果 $\mathfrak{p}$ 和 $\mathfrak{z}$ 是 A 的内与外铺砌, 则

$$\mu(|\mathfrak{p}|) = \int e_{|\mathfrak{p}|} \leqslant \int e_A \leqslant \int e_{|\mathfrak{z}|} = \mu(|\mathfrak{z}|) .$$

因此, $\int e_A$ 在 $\mu^*(A)$ 与 $\overline{\mu}(A)$ 之间, 从而等于 $\mu(A)$. 对任意 $f \in \mathfrak{F}$ 和任意 $A \in \mathfrak{D}_{\min}$, 它们如果满足 $\mathrm{supp}\, f \subset A$, 我们则有 $-\|f\|_\infty e_A \leqslant f \leqslant \|f\|_\infty e_A$, 从而 $|\int f| \leqslant \|f\|_\infty \mu(A)$, 这可由 $\int 3'$ 和 (9.10) 式得到. 在右端对所有这样的集合 A 取下确界便得到了 $\int 5$. □

8.10 可度函数的积分

现在我们按照处理公理 $\mathfrak{D}$ 和 μ 相同的方式来处理公理 $\mathfrak{F}$ 和 $\int$. 我们要构造一个 "极小" 理论而后以逼近方式得到一个 "大" 的理论. 由命题 9.1, 类 $\mathfrak{F}$ 必包含了对任意铺砌 $\mathfrak{p}$ 的函数 $e_{|\mathfrak{p}|}$. 由 $\mathfrak{F}1$ 它从而包含了所有这些函数的线性组合 .

定义 10.1 一个可铺函数是指对某个铺砌 $\mathfrak{p}$, 有一个函数 $f = f_{\mathfrak{p}}$, 其定义为

$$f = \sum_{\square_i \in \mathfrak{p}} c_i e_{\square_i} . \tag{10.1}$$

容易看出, 所有可铺函数的集合满足公理 $\mathfrak{F}1$ 到 $\mathfrak{F}3$. 进一步说, 如果 f 由 (10.1) 式给出, 则从命题 9.1 和公理 $\int 1$ 知道, 积分 $\int$ 在所有可铺函数的类上由

$$\int f = \sum c_i \mu(\square_i) \tag{10.2}$$

惟一决定 .

读者应该验证下面的事实: 如果设 $\mathfrak{F}_P$ 为所有可铺函数的类, $\int$ 由 (10.2) 式给出, 则我们所有的公理都被满足. 别忘了证明 $\int$ 是有明确定义的: 如果 f 以两种方式表达为 (10.1) 式的形式, 则由 (10.2) 式给出的和是相同的 .

显然可铺函数构成了一个太小的函数类. 我们想有一个能包含所有具紧支集的连续函数和所有可度集 A 的特征函数 e_A 的函数类, 例如.

定义 10.2 一个具紧支集的有界函数 f 被称作为可度的是说, 如果对任意 $\varepsilon > 0$ 和 $\delta > 0$, 存在一个可铺函数 $g = g_{\varepsilon,\delta}$ 和一个可度集 $A = A_{\varepsilon,\delta}$, 使得对所有 $\boldsymbol{x} \notin A$ 成立

$$|f(\boldsymbol{x}) - g(\boldsymbol{x})| < \varepsilon \tag{10.3}$$

和

$$\mu(A) < \varepsilon \tag{10.4}$$

称此偶对 $\prec g, A \succ$ 为 f 的一个可铺 ε, δ - 逼近 .

我们来验证所有可度函数的类 $\mathfrak{F}_{\text{con}}$ 满足公理 $\mathfrak{F}1$ 到 $\mathfrak{F}3$. 可以清楚看出, 如果 f 可度, 则对任意常数 a, af 也是可度的. 如果 f_1 和 f_2 可度, 设 $\prec g_1, A_1 \succ$ 和 $\prec g_2, A_2 \succ$ 分别为 f_1 和 f_2 的可铺 ε, δ - 逼近. 于是对所有 $\boldsymbol{x} \notin A_1 \cup A_2$,

$$|(f_1 + f_2)(\boldsymbol{x}) - (g_1 + g_2)(\boldsymbol{x})| < 2\varepsilon$$

以及

$$\mu(A_1 \cup A_2) < 2\delta .$$

因此 $\prec g_1 + g_2, A_1 \cup A_2 \succ$ 给出了 $f_1 + f_2$ 的一个可铺 $2\varepsilon, 2\varepsilon$ - 逼近 .

要证明 $\mathfrak{F}2$, 我们只要看出如果 $\prec g, A \succ$ 是 f 的一个可铺 ε, δ - 逼近, 则 $\prec Tg, TA \succ$ 是 Tf 的一个逼近 .

相似的论证建立了对乘法的类比结果 .

命题 10.1 设 f_1 和 f_2 为两个可度函数, 则 $f_1 f_2$ 也可度.

证明 设 M 对所有 $\boldsymbol{x}$ 满足 $|f_1(\boldsymbol{x})| < M, |f_2(\boldsymbol{x})| < M$. 回想两个可铺函数的乘积仍是可铺的. 用前面相同的记号, 我们对所有 $\boldsymbol{x} \notin A_1 \cup A_2$ 有

$$\begin{aligned}|f_1 f_2(\boldsymbol{x}) - g_1(\boldsymbol{x}) g_2(\boldsymbol{x})| &\leqslant |f_1(\boldsymbol{x})|\ |f_2(\boldsymbol{x}) - g_2(\boldsymbol{x})| + |g_2(\boldsymbol{x})|\ |f_1(\boldsymbol{x}) - g_1(\boldsymbol{x})| \\ &< M\varepsilon + (M + \varepsilon)\varepsilon .\end{aligned}$$

因此 $\prec g_1 g_2, A_1 \cup A_2 \succ$ 是 $f_1 f_2$ 的可铺 $(2M + \varepsilon)\varepsilon, 2\delta$ - 逼近. □

至于 $\mathfrak{F}3$, 立即知道一个更强的形式成立：

命题 10.2 *如果 B 是可度集则 e_B 是个可度函数.*

证明 事实上, 设 $\mathfrak{p}$ 是 B 的一个内铺砌, 满足 $\mu(B)-\mu(|\mathfrak{p}|)<\delta$. 于是如果 $x\notin B-|\mathfrak{p}|$ 便有

$$e_B(x)-e_{|\mathfrak{p}|}(x)=0\ ,$$

和

$$\mu(B-|\mathfrak{p}|)<\delta\ ,$$

故 $\prec e_{|\mathfrak{p}|},|\mathfrak{p}|\succ$ 是 e_B 对任意 $\varepsilon>0$ 的一个 ε,δ - 逼近. □

我们现在建立对可度函数的另一种特殊描述.

命题 10.3 *函数 f 可度的充要条件是对任意的 ε, 存在可铺函数 h 和 k, 使得 $h\leqslant f\leqslant k$, 且 $\int(k-h)<\varepsilon$.*

证明 如果 f 是可度的, 设 R 为包含 $\operatorname{supp} f$ 的一个矩形. 设 $\prec g,A\succ$ 为 f 的一个 ε,δ - 逼近. 又设 P 为包含 $A=A_{\varepsilon,\delta}$ 的一个可铺集使得 $\mu(P)<\delta$, 并设 m 为 $|f|$ 的一个界. 于是 $g-\varepsilon e_R-me_P\leqslant f\leqslant g+\varepsilon e_R+me_P$, 其中两端的函数显然是可铺的, 而它们的积分的差小于 $2\varepsilon\mu(R)+2m\delta$. 由于 ε,δ 任意, 我们便找出了所要的 h 和 k. 反之, 如果 h 和 k 为可铺函数, 使得 $h\leqslant f\leqslant k$ 和 $\int(k-h)<a$, 于是使 $k-h\geqslant a^{\frac{1}{2}}$ 的点集是一个可铺集 A. 另外, $a^{\frac{1}{2}}\mu(A)\leqslant\int e_A(k-h)\leqslant\int(k-h)\leqslant a$, 故 $\mu(A)\leqslant a^{\frac{1}{2}}$. 给出 ε 和 δ, 我们只需选取 $a\leqslant\min(\varepsilon^2,\delta^2)$, 并取 g 或者为 k 或者为 h 便可看出 f 是可度的. □

推论 10.1 *如果对每个 ε 存在可度函数 f_1 和 f_2 使得 $f_1\leqslant f\leqslant f_2$ 和 $\int(f_2-f_1)<\varepsilon$, 则函数 f 可度.*

证明 因为如此一来, 我们便可以找到可铺函数 $h\leqslant f_1$ 和 $k\geqslant f_2$, 使得 $\int(f_1-h)<\varepsilon$ 和 $\int(k-f_2)<\varepsilon$, 最后便有了 $h\leqslant f\leqslant k$ 且 $\int(k-h)<3\varepsilon$. □

定理 10.1 *设 $\mathfrak{F}$ 是个满足公理 $\mathfrak{F}1$ 到 $\mathfrak{F}3$ 的函数类, 使得 $\mathfrak{F}_p\subset\mathfrak{F}\subset\mathfrak{F}_{\text{con}}$. 于是存在 $\mathfrak{F}$ 上惟一的 $\int$ 满足从 $\int 1$ 到 $\int 4$ 的公理.*

证明 如果 $\int$ 是 $\mathfrak{F}$ 上满足公理 $\int 1$ 到 $\int 4$ 的任一积分, 我们则必定同时有对可铺的 $h\leqslant f$, $\int f$ 等于 $\operatorname{lub}\int h$, 又对 $\geqslant f$ 的可铺 k 它等于 $\operatorname{glb}\int k$; 这可由命题

10.3 推出. 因此这积分在 $\mathfrak{F}$ 上惟一地确定. 另外容易看出, 如果此积分在 $\mathfrak{F}$ 上由 $\int f = \text{lub} \int h = \text{glb} \int k$ 定义, 由于对可铺函数上惟一确定的积分而言公理 $\int 1$ 到 $\int 4$ 成立, 所以这些公理对 $\mathfrak{F}$ 也成立. □

习题

10.1 设 f 和 g 为可度函数, 它们对 $x \notin A$ 有 $f(x) = g(x)$, 其中 $\mu(A) = 0$. 于是 $\int f = \int g$. (这表明就积分这个目的而言, 我们只需要在容度零的集合程度上了解一个函数.)

定义 10.3 设 f 为可度函数, A 为一可度集. 我们称 $\int e_A f$ 为 f 在 A 上的积分并记为 $\int_A f$. 因此

$$\int_A f = \int e_A f \,. \tag{10.5}$$

公理 $\int 1$ 和 (9.7) 式的一个直接推论是

$$\int_{A_1 \cup A_2} f = \int_{A_1} f + \int_{A_2} f - \int_{A_1 \cap A_2} f \,. \tag{10.6}$$

习题 10.1 的一个直接推论是

$$\left| \int_A f \right| \leqslant \sup_{x \in A} |f(x)| \mu(A)$$

我们现在给出可度函数的另一个有用的特征描述来结束本节.

命题 10.4 设 f 为具紧支集的一个有界函数. 于是 f 可度的充要条件是对每个 $\varepsilon > 0$ 和 $\delta > 0$, 我们可找到一个 $\eta > 0$ 和一个可度集 A_δ, 使得 $\mu(A_\delta) < \delta$, 以及只要 $\|x - y\| < \eta$ 和 $x, y \notin A_\delta$ 便有

$$|f(x) - f(y)| < \varepsilon \,. \tag{10.7}$$

证明 假定对每个 ε, δ 我们可以找到这样的 η 和 A_δ. 设 $\mathfrak{p} = \{\square_i\}$ 为一个铺砌, 使得

(i) $\text{supp}\, f \subset |\mathfrak{p}|$;

(ii) 如果 $x, y \in \square_i$, 则 $\|x - y\| < \eta$;

(iii) 如果 $\mathfrak{z} = \{\square_i \in \mathfrak{p} : \square_i \cap A_\delta \neq \varnothing\}$, 则 $\mu(|\mathfrak{z}|) < 2\delta$. 然后当 $x \in \square_i$ 时, 设 $f_{\varepsilon,2\delta}(x) = f(x_i)$, 其中 x_i 为 $\square_i$ 中某个点. 由 (ii) 和 (iii) 我们看出 $\prec f_{\varepsilon,2\delta}, |\mathfrak{z}| \succ$ 是 f 的一个可铺 $\varepsilon, 2\delta$ - 逼近. 因此 f 是可度的.

反之, 假设 f 是可度的, 并设 $\prec f_{\varepsilon/2,\delta/2}, A_{\varepsilon/2,\delta/2} \succ$ 为 f 的一个可铺逼近.

设 $\mathfrak{p}=\{\square_i\}$ 为 $f_{\varepsilon/2,\delta/2}$ 所相伴的铺砌. 将 $\square_i$ 依其中心按比例因子 $(1-\xi)$ 收缩得到矩形 $\square_i'$(见图 8.10), 然后用矩形 $\square_i'$ 取代 $\square_i$. 因此 $\mu(\square_i')=(1-\xi)^n\mu(\square_i)$. 对任意 $\boldsymbol{x},\boldsymbol{y}\in\cup\square_i'$, 如果

$$\|\boldsymbol{x}-\boldsymbol{y}\|<\eta\,,$$

其中 η 充分小, 则 $\boldsymbol{x},\boldsymbol{y}$ 属于同一 $\square_i'$. 如果

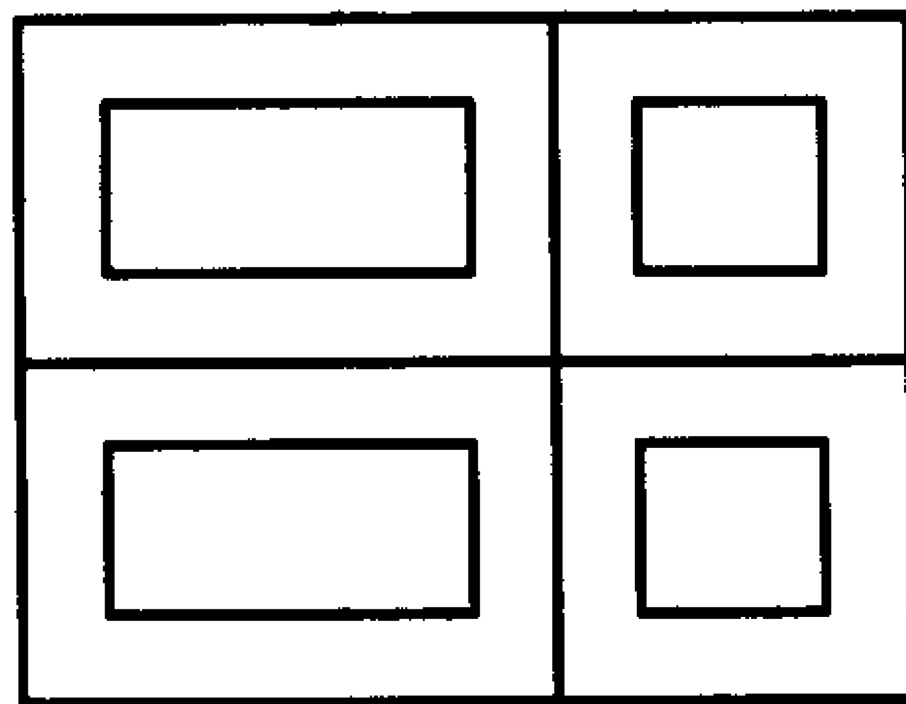


图 8.10

$$\boldsymbol{x},\boldsymbol{y}\in\cup\square_i',\quad x,y\notin A_{\varepsilon/2,\delta/2}\quad 和\quad \|\boldsymbol{x}-\boldsymbol{y}\|<\eta\,,$$

则

$$|f(\boldsymbol{x})-f(\boldsymbol{y})|\leqslant|f(\boldsymbol{x})-f_{\varepsilon/2,\delta/2}(\boldsymbol{x})|+|f(\boldsymbol{y})-f_{\varepsilon/2,\delta/2}(\boldsymbol{y})|$$
$$+|f_{\varepsilon/2,\delta/2}(\boldsymbol{x})-f_{\varepsilon/2,\delta/2}(\boldsymbol{y})|\,.$$

但第三项为零, 故 $|f(\boldsymbol{x})-f(\boldsymbol{y})|<\varepsilon$. 由于首先选取 ξ 充分小, 我们可安排使 $\mu(|\mathfrak{p}|-\cup\square_i')<\delta/2$. 然后我们可以选取 δ 如此之小, 使得 $\|\boldsymbol{x}-\boldsymbol{y}\|<\eta$ 蕴涵了如果 $\boldsymbol{x},\boldsymbol{y}\in\cup\square_i'$ 则 x,y 属于同一个 $\square_i'$. 对这个 η 和对 $A_\delta=A_{\varepsilon/2,\delta/2}\cup(|\mathfrak{p}|-\cup\square_i')$, 等式 (10.7) 成立, 而且 $\mu(A_\delta)<\delta$. □

特别地, 在一个零容度集之外连续并且有紧支集的一个有界函数是可度的 .

习题

10.2 证明对任何有界集 A, e_A 是可度函数当且仅当 A 是个可度集 .

10.3 设 f 是个可度函数, 其支集含于立方体 $\square$ 内. 对每个 δ, 设 $\mathfrak{p}_\delta=\{\square_{i,\delta}\}_{i\in I_\delta}$ 为一个铺砌, 使 $|\mathfrak{p}_\delta|=\square$, 而且其中每个立方体的直径小于 δ. 设 $\boldsymbol{x}_{i,\delta}$ 为 $\square_{i,\delta}$ 中某个点. 称表达式

$$\sum_{i\in I_\delta}f(\boldsymbol{x}_{i,\delta})\mu(\square_{i,\delta})$$

为 f 的一个黎曼 δ - 近似和. 证明对任意 $\varepsilon>0$, 存在一个 $\delta_0[=\delta_0(f)]>0$, 使得只要 $\delta<\delta_0$. 便有

$$\left|\int f-\sum_{i\in I_\delta}f(\boldsymbol{x}_{i,\delta})\mu(\square_{i,\delta})\right|<\varepsilon\,.$$

8.11 换元公式

本节致力于下面定理的证明; 此定理具有基本的重要性.

定理 11.1 设 U,V 为 $\mathbb{R}^n$ 中的有界开集, φ 是由 U 到 V 上的一个连续可微的一一映射, 其逆 φ^{-1} 可微. 又设 f 为一可度函数, 其支集 $\operatorname{supp} f \subset V$. 那么, $(f \circ \varphi)$ 是个可度函数, 且

$$\int_V f = \int_U (f \circ \varphi)|\det J_\varphi| . \tag{11.1}$$

回想起如果映射 φ 由 $y^i = \varphi^i(x^1, \cdots, x^n)$ 给出, 则 J_φ 是个矩阵为 $[\partial\varphi_i/\partial x_j]$ 的线性变换.

又注意到如果 φ 是个非异线性变换 (因而 J_φ 恰是 φ), 则定理 11.1 是定理 8.1 的一个简单推论. 事实上, 对形如 e_A 的函数我们看出有 $e_A \circ \varphi = e_{\varphi^{-1}(A)}$, 而且等式 (11.1) 在这种简单情形约化成了 (8.1) 式. 由线性性, (11.1) 式对所有可铺函数成立.

另外, $f \circ \varphi$ 是可度的. 假设 $|f(\boldsymbol{x}) - f(\boldsymbol{y})|$ 在 $\|\boldsymbol{x} - \boldsymbol{y}\| < \|\varphi\|$, 其中 $\boldsymbol{x}, \boldsymbol{y} \notin A$ 而 $\mu(A) < \delta$ 时小于 ε. 于是当

$$\|\boldsymbol{u} - \boldsymbol{v}\| < \eta/\|\varphi\| \quad 和 \quad \boldsymbol{u}, \boldsymbol{v} \notin \varphi^{-1}(A) ,$$

其中 $\mu|\varphi^{-1}(A)| < \delta/|\operatorname{def} \varphi|$ 时 $|f \circ \varphi(\boldsymbol{u}) - f \circ \varphi(\boldsymbol{v})| < \varepsilon$.

现设 $g_{\varepsilon,\delta}, A_{\varepsilon,\delta}$ 是 f 的一族逼近的可铺函数. 于是对 $\boldsymbol{x} \notin \varphi^{-1}(A_{\varepsilon,\delta})$ 而 $\mu(\varphi^{-1}A_{\varepsilon,\delta}) < \delta/|\det \varphi|$ 有 $|f \circ \varphi(\boldsymbol{x}) - g_{\varepsilon,\delta} \circ \varphi(\boldsymbol{x})| < \varepsilon$. 因此 $\int (g_{\varepsilon,\delta} \circ \varphi)|\det \varphi| \to \int (f \circ \varphi)|\det \varphi|$, 从而等式 (11.1) 对所有可度函数 f 成立.

对于非线性映射的定理 11.1 的证明更多了一点技巧性. 它实质上是用线性映射局部地逼近 φ 构成, 我们将分成若干步来完成它. 我们在 $\mathbb{R}^n$ 上要使用一致范数 $\|\boldsymbol{x}\|_\infty = \max|x^i|$. 因为一个球在此范数下实际上是个立方体, 故较为方便, 但这种细节上的优美不是非要不可的.

设 ψ 为一 (连续) 可微映射, 它定义在一个凸开集 U 上. 如果立方体 $\square = \square_{\boldsymbol{p}-r\boldsymbol{1}}^{\boldsymbol{p}+r\boldsymbol{1}} \subset U$, 则中值定理 (第三章第 7 节) 表明对任意 $y \in \square$,

$$\|\psi(\boldsymbol{y}) - \psi(\boldsymbol{p})\|_\infty < \|\boldsymbol{y} - \boldsymbol{p}\|_\infty \sup_{\boldsymbol{z}\in\square} \|J_\psi(\boldsymbol{z})\| .$$

因此

$$\psi(\square) \subset \square_{\psi(\boldsymbol{p})-Kr\boldsymbol{1}}^{\psi(\boldsymbol{p})+Kr\boldsymbol{1}}, \quad 其中 K = \sup_{\boldsymbol{z}\in\square} \|J_\psi(\boldsymbol{z})\| .$$

于是

$$\mu(\psi(\square)) < (\sup_{z\in\square}\|J_\psi(z)\|)^n\mu(\square) . \tag{11.2}$$

引理 11.1 设 φ 同于定理 11.1 中的 φ. 于是对任何 $\overline{A}\subset U$ 的可度集 A 有

$$\mu(\varphi(A)) \leqslant \int_A |\det J_\varphi| . \tag{11.3}$$

证明 将公式 (11.2) 用于映射 $\psi = L^{-1}\varphi$, 其中 L 是线性变换. 于是

$$|\det L|^{-1}\mu(\varphi(\square)) = \mu(L^{-1}\varphi(\square)) \leqslant (\sup_{z\in\square}\|J_{L^{-1}\varphi}(z)\|)^n\mu(\square) .$$

由于 $J_{L^{-1}\varphi} = L^{-1}J_\varphi$, 我们得到

$$\mu(\varphi(\square)) \leqslant |\det L|(\sup_{z\in\square}\|L^{-1}J_\varphi(z)\|^n\mu(\square) , \tag{11.4}$$

其中 $\square$ 为任意包含于 φ 的定义域中的立方体, 而 L 为任意线性变换 .

对任意 $\varepsilon > 0$, 设 δ 如此之小使得对所有在 A 的一个紧邻域中的 x,y 有 $\|x-y\|_\infty < \delta$ 时, 成立 $\|J_\varphi(x)^{-1}J_\varphi(y)\| < 1+\varepsilon$. (因为 $J(x)$ 是 x 的一致连续函数, 从而当 x 靠近 y 时 $J_\varphi(x)^{-1}J_\varphi(y)$ 靠近单位矩阵, 那么选取这样的 δ 是可能的. 见第四章第 8 节 .)

选一个 A 的外铺砌 $\mathfrak{z} = \{\square_i\}$, 其中 $\square_i$ 是所有边长小于 δ 的立方体. 设 x_i 为 $\square_i$ 中的点. 于是把 (11.4) 式用于每个 $\square_i$, 而取 $L = J_\varphi(x_i)$, 我们得到

$$\mu(\varphi(A)) < \mu(\varphi(|\mathfrak{z}|)) = \sum \mu\varphi(\square_i) < \sum |\det J_\varphi(x_i)|(1+\varepsilon)^n\mu(\square_i) .$$

我们也可假定取 δ 足够小使得对所有 $z\in\square_i$ 及所有 i ,

$$|\det J_\varphi(z)| > (1-\varepsilon)|\det J_\varphi(x_i)| .$$

于是我们有

$$\int_{\square_i} |\det J_\varphi| > (1-\varepsilon)|\det J_\varphi(x_i)|\mu(\square_i) ,$$

故而

$$\mu(\varphi(A)) < \frac{1}{1-\varepsilon}(1+\varepsilon)^n\int_{|\mathfrak{z}|} |\det J_\varphi| .$$

因为 ε 为任意, $\mathfrak{z}$ 为 A 的任意外铺砌, 故而得到了 (11.3) 式 . □

我们现在可以得出结论说对任何 $\operatorname{supp} f \subset V$ 的可度的 $f, f\circ\varphi$ 也是可度的. 事实上, 设 K 可选得如此之大, 使其为 φ 在 $\varphi^{-1}(\operatorname{supp} f)$ 上的一个利普希茨常数, 并使 $K > |\det J_{\varphi^{-1}}(\boldsymbol{u})|$, 其中 $\boldsymbol{u} \in \operatorname{supp} f$. 对给出的 ε 和 δ, 我们现在能够找到一个 η, 使得如果 $\|\boldsymbol{u}-\boldsymbol{v}\| < \eta, \boldsymbol{u}, \boldsymbol{v} \notin A_\delta$ 而 $\mu(A_\delta) < \delta$, 有

$$|f(\boldsymbol{u}) - f(\boldsymbol{v})| < \varepsilon .$$

但这意味着如果 $\|\boldsymbol{x} - \boldsymbol{y}\| < \eta/K$ 和 $\boldsymbol{x}, \boldsymbol{y} \notin \varphi^{-1}(A\delta)$, 且 $\mu(\varphi^{-1}(A\delta)) < K_\delta$, 由 (11.3) 式有

$$|f\circ\varphi(\boldsymbol{x}) - f\circ\varphi(\boldsymbol{y})| < \varepsilon .$$

因为 K 的选取与 ε 和 δ 无关, 这表明 $f\circ\varphi$ 是可度的 .

引理 11.2 设 φ, U 和 V 同于定理 11.1 中的. 设 f 为非负可度函数, 满足 $\operatorname{supp} f \subset V$. 于是

$$\int_V f \leqslant \int_U (f\circ\varphi)|\det J_\varphi| . \tag{11.5}$$

证明 设 $\prec g, A \succ$ 为 f 的一个可铺 ε, δ - 逼近, 它对所有 $\boldsymbol{u}$ 满足 $g(\boldsymbol{u}) \leqslant f(\boldsymbol{u})$. 如果 $\mathfrak{p} = \{\square_i\}$ 为与 g 相伴的铺砌, 我们可假定有 $\operatorname{supp} f \subset |\mathfrak{p}|$. 于是,

$$\begin{aligned}\int g &= \sum_{\substack{i \\ u_i \in \square_i}} g(\boldsymbol{u}_i)\mu(\square_i) \leqslant \sum g(\boldsymbol{u}_i) \int_{\varphi^{-1}(\square_i)} |\det J_\varphi| \\ &\leqslant \sum_i \int_{\varphi^{-1}(\square_i)} (f\circ\varphi)|\det J_\varphi| \\ &= \int_{\cup\varphi^{-1}(\square_i)} (f\circ\varphi)|\det J_\varphi| = \int (f\circ\varphi)|\det J_\varphi| .\end{aligned}$$

因为我们可选取 g 使得 $\int g \to \int f$, 便得到了 (11.5) 式 . □

引理 11.3 设 φ, U, V 和 f 如同定理 11.1 中的. 设 f 为一非负函数. 于是等式 (11.1) 成立 .

证明 对映射 φ^{-1} 和函数 $(f\circ\varphi)|\det J_\varphi|$ 运用 (11.5) 式. 因为 $J_\varphi(\boldsymbol{x}) \circ J_{\varphi^{-1}}(\varphi(\boldsymbol{x})) = \mathrm{id}$, 我们得到

$$\begin{aligned}\int (f\circ\varphi)|\det J_\varphi| &\leqslant \int [(f\circ\varphi)\circ\varphi^{-1}](|\det J_\varphi|\circ\varphi^{-1})|\det J_{\varphi^{-1}}| \\ &= \int f .\end{aligned}$$

将此与 (11.5) 式结合便证明了此引理 . □

做完定理 11.1 的证明 任意实的可度函数可以写成两个正的可度函数的差. 如果对所有 x, 对某个大的数 M 有 $f(x) > -M$, 我们可写出 $f = (f + Me_{\square}) - Me_{\square}$, 其中 $\operatorname{supp} f \subset \square$. 因为我们已对非负函数验证了等式 (11.1), 并且由于 (11.1) 式的两端对 f 是线性的, 故已证好. 相似地, 任何有界的复可度函数 f 可写为 $f = f_1 + if_2$, 其中 f_1 和 f_2 为有界实可度函数, 故也得证 . □

在实际使用中, 我们有时在从严格定义上说定理 11.1 的假设条件在没有验证的情形下便使用了等式 (11.1). 例如, 在 $\mathbb{R}^2$ 中我们可能需要引进 "极坐标". 就是说, 设 r,θ 为 $\mathbb{R}^2$ 上的坐标; 如果 S 是集合 $0 \leqslant \theta < 2\pi, 0 \leqslant r$, 我们可以考虑 $\varphi : S \to \mathbb{R}^2$, 它由 $x = r\cos\theta, y = r\sin\theta$ 给出, 其中 x,y 为 $\mathbb{R}^2$ 上的第二个拷贝上的坐标. 这个映射在 $r > 0$ 上是一一的并具有正的雅可比. 如果我们考虑开集 $U \subset S : 0 < r, 0 < \theta \leqslant 2\pi$ 和 $V \subset \mathbb{R}^2$, 其中 $V = \mathbb{R}^2 - \{x,y : y = 0, x \geqslant 0\}$, 则定理 11.1 的假设条件是满足的, 从而如果 $\operatorname{supp} f \subset V$ 我们可以写出 (因为 $\det J_{\varphi} = r$)

$$\int f = \int (f \circ \varphi) r \,. \tag{11.6}$$

但等式 (11.6) 在没有 $\operatorname{supp} f \subset V$ 的限制下也是成立的. 事实上, 如果 D_ε 是一条宽为 ε 的围绕射线 $y = 0, x \geqslant 0$ 的长条, 则 $f = fe_{D_\varepsilon} + fe_{\mathbb{R}^n - D_\varepsilon}$, 而且当 $\varepsilon \to 0$ 时 $\int fe_{D_\varepsilon} \to 0$(图 8.11). 相似地, $\int (f \circ \varphi)(r \circ \varphi) e_{D_\varepsilon} \circ \varphi \to 0$, 故由简单的极限式论证知 (11.6) 式对所有的可度 f 成立 .

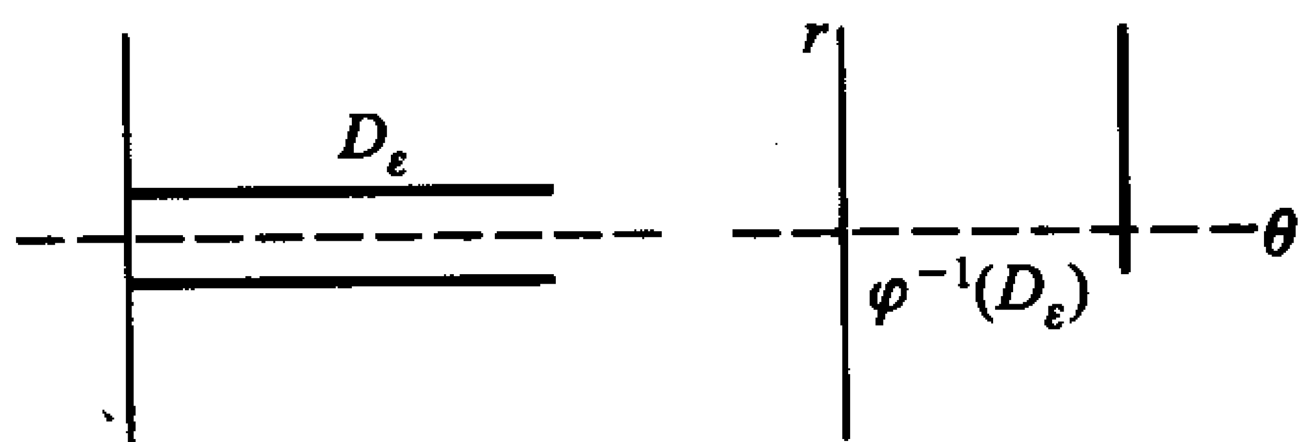


图 8.11

我们不再讲包括了定理 11.1 的所有这类有用扩充的一般性定理. 在每一种情形中, 极限式的论证常常是相当直接的, 从而留给读者自己去做 .

习题

11.1 由 $\boldsymbol{v}^1, \cdots, \boldsymbol{v}^n$ 张成的平行六面体指的是所有 $x = \sum_1^n \xi^i \boldsymbol{v}^i$ 的集合, 其中 $0 \leqslant \xi^i < 1$. 证明此平行六面体的容度由 $|\det((\boldsymbol{v}_i, \boldsymbol{v}_j))|^{\frac{1}{2}}$ 给出 .

11.2 利用单位球的容度来表达椭球

$$\left\{x : \frac{(x^1)^2}{(a^1)^2} + \cdots + \frac{(x^n)^2}{(a^n)^2} \leqslant 1\right\}$$

的容度 .

11.3 计算映射 $\prec r, \theta \succ \mapsto \prec x, y \succ$ 的雅可比行列式, 其中 $x = r\cos\theta$, $y = r\sin\theta$.

11.4 计算映射 $\prec r, \theta, \varphi \succ \mapsto \prec x, y, z \succ$ 的雅可比行列式, 其中 $x = r\cos\varphi\sin\theta$, $y = r\sin\varphi\cos\theta$, $z = r\cos\theta$.

11.5 计算映射 $\prec r, \theta, z \succ \mapsto \prec x, y, z \succ$ 的雅可比行列式, 其中 $x = r\cos\theta$, $y = r\sin\theta$, $z = 2$.

8.12 累次积分

在单变量的情形, 即 $\mathbb{R}^1$ 上的积分理论, 微积分基本定理把一个函数的积分计算化成它的反导数的计算. 此定理在 n 维的推广将在后面的章节中讲述. 在本节中我们将证明, 在许多情形中一个 n 维的积分可以化为 n 个累次的一维积分 .

假设我们以一种固定的方式将 $\mathbb{R}^n$ 看作为直积 $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$. 我们将把每个 $z \in \mathbb{R}^n$ 写成 $z = \prec x, y \succ$, 其中 $x \in \mathbb{R}^k, y \in \mathbb{R}^l$.

定义 12.1 *我们说一个可度函数 f 是相对于分解 $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$ 可度指的是如果存在一个集合 $A_f \subset \mathbb{R}^k$ 是在 $\mathbb{R}^k$ 中容度为 0, 使得*

(i) 对每个固定的 $x \in \mathbb{R}^k, x \notin A_f$, 函数 $f(x, \cdot)$ 是 $\mathbb{R}^l$ 上的可度函数 ;

(ii) 函数 $\int_{\mathbb{R}^l} f$ 对每个 x 给定了数 $\int_{\mathbb{R}^l} f(x, \cdot)$ 是 $\mathbb{R}^k$ 上的可度函数 .

容易看出, 所有这些函数的集合满足公理 $\mathfrak{F}1$ 到 $\mathfrak{F}3$. (不能立刻得出的只有公理 $\mathfrak{F}2$. 但这只是下述事实的一个简单推论: 任何一个平移 T 可以重写为 T_1T_2, 其中 T_1 是 $\mathbb{R}^k$ 中的平移, 而 T_2 是 $\mathbb{R}^l$ 中的平移 .)

同样容易验证把任意这样的 f 指定数

$$\int_{\mathbb{R}^k} \left(\int_{\mathbb{R}^l} f(\cdot, \cdot) \right)$$

的这个规则满足公理 $\int 1$ 到 $\int 4$. 不是立刻显见的只有公理 $\int 3$. 然而, 如果 $\mathfrak{p}$ 为 $\operatorname{supp} f \subset |\mathfrak{p}|$ 的任意的铺砌, 则

$$f \leqslant \|f\| e_{|\mathfrak{p}|}$$

和

$$\int_{\mathbb{R}^k}\left(\int_{\mathbb{R}^l}\|f\|e_{|\mathfrak{p}|}\right)=\|f\|\int_{\mathbb{R}^k}\int_{\mathbb{R}^l}e_{|\mathfrak{p}|}=\|f\|\mu(e_{|\mathfrak{p}|})\ ,$$

这是因为对任何矩形都有

$$\int_{\mathbb{R}^k}\int_{\mathbb{R}^l}e_{\square}=\mu(\square)$$

(直接验证). 因此, 由定理 10.1 中的惟一性部分得出

$$\int_{\mathbb{R}^n}f=\int_{\mathbb{R}^k}\left(\int_{\mathbb{R}^l}f(\cdot,\cdot)\right)\ . \tag{12.1}$$

特别留意一下, 如果 f 也相对于分解 $\mathbb{R}^n=\mathbb{R}^l\times\mathbb{R}^k$ 也是可度的, 则

$$\int_{\mathbb{R}^n}f=\int_{\mathbb{R}^l}\left(\int_{\mathbb{R}^k}f(\cdot,\cdot)\right)\ .$$

于是特别地, 对这样的 f 的二重积分与顺序无关 .

在实际应用中, 我们将遇到的所有函数都是相对于 $\mathbb{R}^n$ 的任意分解可度的. 特别写 $\mathbb{R}^n=\mathbb{R}^1\times\cdots\times\mathbb{R}^1$, 我们有

$$\int_{\mathbb{R}^n}f=\int(\cdots(f(\cdot,\cdots,\cdot)))\ . \tag{12.2}$$

利用直角坐标 $x^1,\cdots,x^n$, 上面最后一个表达式通常被写成

$$\int\left(\cdots\left(\int f(x^1,\cdots,x^n)dx^1\right)\cdots\right)dx_n\ .$$

因为这个原因, (12.2) 式左端的表达式常常写为

$$\int_{\mathbb{R}^n}f=\int\cdots\int f(x^1,\cdots,x^n)dx^1\cdots dx^n\ .$$

让我们来做些简单的例题, 以解释前面几节所给出的各种积分方法 .

例题 12.1 计算顶角 α 的实锥体 (顶点在 O) 与球壳 $1\leqslant r\leqslant 2$ 相交部分的体积 (图 8.12). 在某个欧几里得运动下不妨设锥的轴为 z 轴. 如果引进极坐标, 我们看出所讨论的集合是集合

$$\square_{\prec 1,0,0\succ}^{\prec 2,2\pi,\alpha/2\succ},\quad 1\leqslant r<2,\quad 0\leqslant\varphi<2\pi,\quad 0\leqslant\theta\leqslant\alpha/2$$

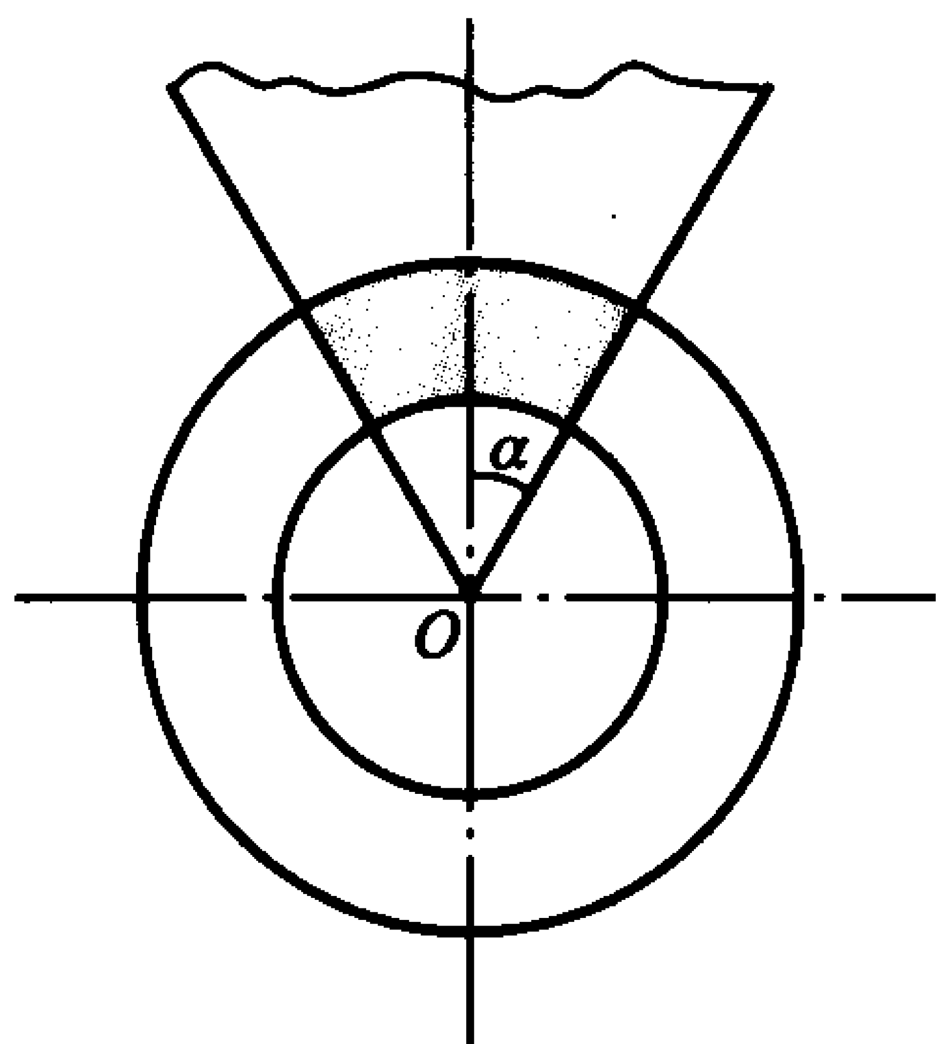


图 8.12

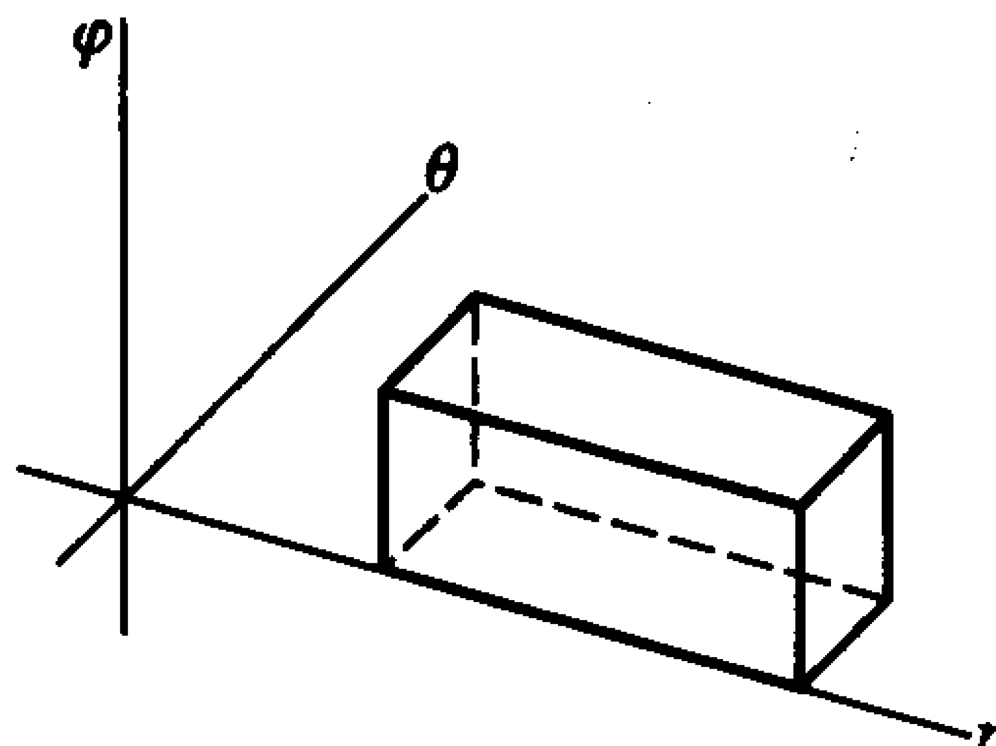


图 8.13

在 $\prec r,\varphi,\theta \succ$ 空间中的像 (图 8.13) . 由换元公式和习题 11.4 知道所讨论的体积由下式给出

$$\begin{aligned}\int r^2\sin\theta &= \int_1^2\int_0^{2\pi}\int_0^{\alpha/2} r^2\sin\theta d\theta d\varphi dr\\ &= 2\pi\int_1^2\int_0^{\alpha/2} r^2\sin\theta d\theta dr\\ &= 2\pi\int_1^2 [1-\cos\alpha/2]r^2dr\\ &= 2\pi[1-\cos(\alpha/2)]\left(\frac{8}{3}-\frac{1}{3}\right).\end{aligned}$$

例题 12.2 设 B 是平面中一个可度集, 而 f_1 和 f_2 为定义在 B 上的两个可度函数. 设 A 为所有 $\prec x,y,z\succ\in\mathbb{E}^3$ 中使 $\prec x,y\succ\in B$ 和 $f_1(x,y)\leqslant z\leqslant f_2(x,y)$ 的集合. 如果 G 为 A 上的任意可度函数, 我们则可表达积分 $\int_A G$ 为

$$\int_A G=\int_A\left\{\int_{f_1(x,y)}^{f_2(x,y)} G(x,y,z)dz\right\}dxdy\ .$$

例如,计算积分 $\int_A z$, 其中 A 为在曲面 $z=x^2+y^2$ 上面的单位球中所有点的集合(图 8.14).

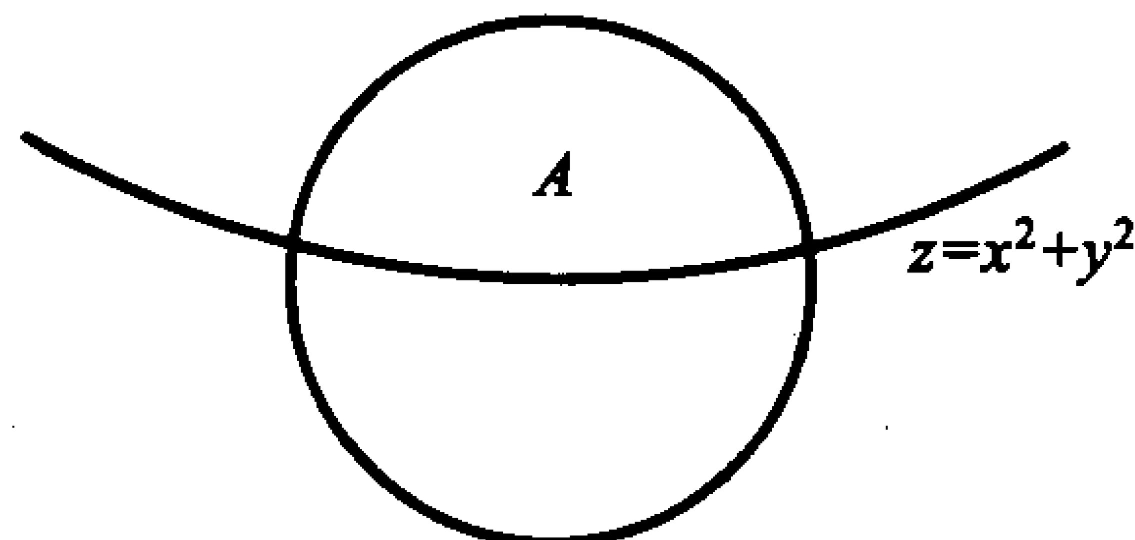


图 8.14

因此

$$A=\{\prec x,y,z\succ: x^2+y^2+z^2\leqslant 1, z\geqslant x^2+y^2\}\ .$$

我们必定有 $x^2+y^2 \leqslant a$, 其中 $a^2+a=1$[故 $a=(\sqrt{5}-1)/2$], 从而能使 $\prec x,y,z \succ$ 属于 A. 于是 $f_1(x,y)=x^2+y^2, f_2(x,y)=\sqrt{1-(x^2+y^2)}$, 而

$$\int_{f_1(x,y)}^{f_2(x,y)} zdz = \frac{1}{2}[1-(x^2+y^2)-(x^2+y^2)^2] ,$$

故利用平面中的极坐标 (和习题 11.3) 有

$$\int_A z = \frac{1}{2}\int_{x^2+y^2\leqslant a}[1-(x^2+y^2)-(x^2+y^2)^2] = \pi\int_r^{\sqrt{a}} r(1-r^2-r^4)dr .$$

正如在最后这个例题中看到的, 作为累次积分来计算一个积分的部分问题是用向量空间分解的手段来确定对积分区域的一个好的描述. 通常用画出图形使情形形像化会有极大的帮助 .

例题 12.3 (计算一个旋转曲面所包围的体积) 这里我们已知一个单变函数 f, 考虑绕 z 轴旋转曲线 $x=f(z), z_1 \leqslant z \leqslant z_2$ 所得到的曲面 (图 8.15). 我们希望计算 $\mu(A)$, 其中

$$A=\{\prec x,y,z \succ : \sqrt{x^2+y^2} \leqslant f(z), z_1 \leqslant z \leqslant z_2\} .$$

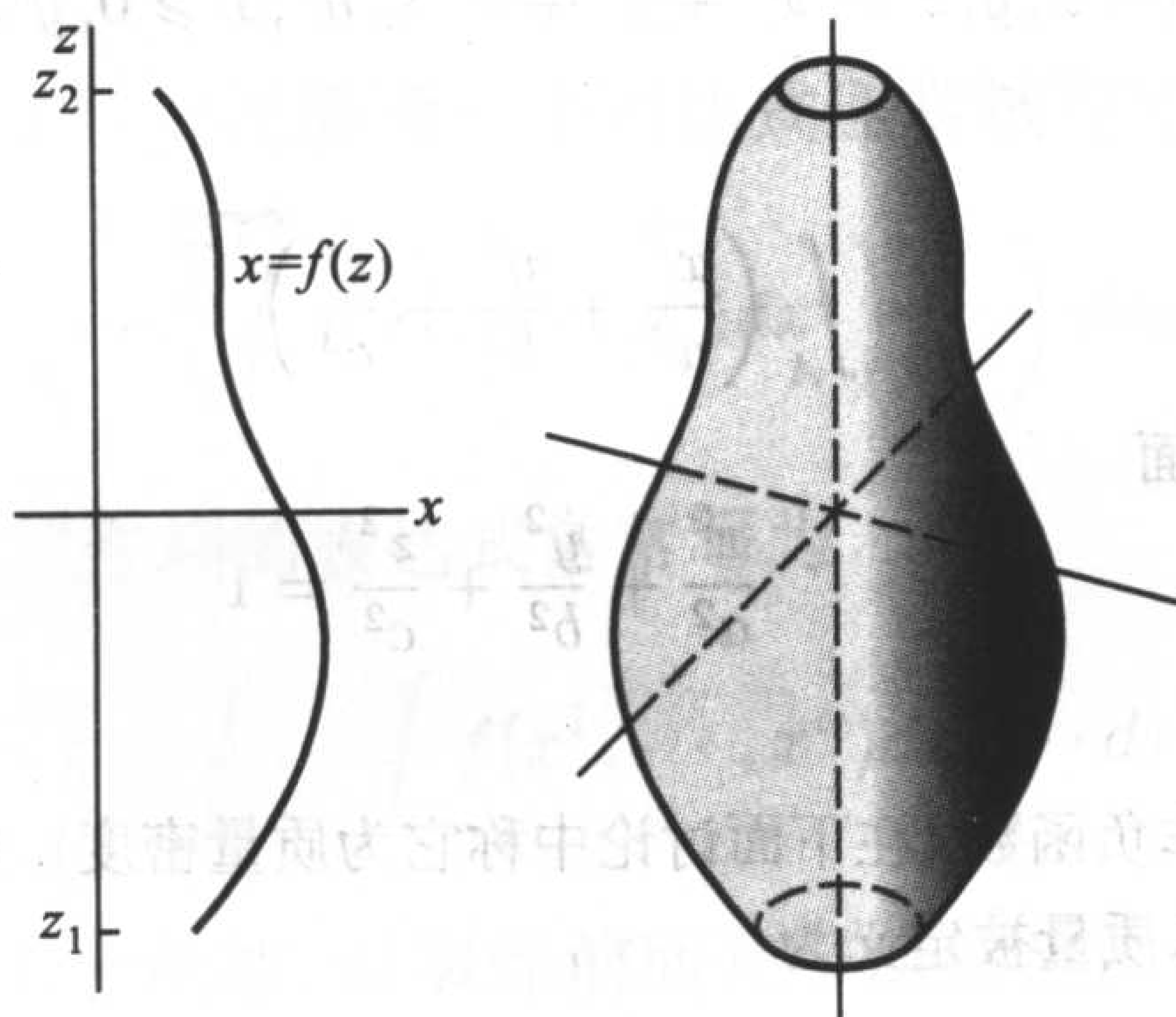


图 8.15

显然在这里使用柱坐标是方便的, 那么我们看出 A 是下面集合 B 的像, 这里的 B 在 $\prec r,\theta,z \succ$ 空间中 :

$$B=\{\prec r,\theta,z \succ : r \leqslant f(z), 0 \leqslant \theta < 2\pi, z_1 \leqslant z \leqslant z_2\} .$$

由习题 11.5, 我们要计算

$$\int_B r = \int_0^{2\pi}\int_{z_1}^{z_2}\int_0^{f(z)} r dr dz d\theta = 2\pi\int_{z_1}^{z_2}\left(\int_0^{f(z)} r dr\right)dz$$
$$= 2\pi\int_{z_1}^{z_2}\frac{f(z)^2}{2}dz .$$

因此

$$\mu(A) = \pi\int_{z_1}^{z_2} f(z)^2 dz .$$

习题

12.1 计算曲面 $z = x^2 + y^2$ 和 $z = x + y$ 之间的体积 .

12.2 求 $\mathbb{E}^3$ 中由平面 $z = 0$, 柱面 $x^2 + y^2 = 2x$ 和锥 $z = +\sqrt{x^2+y^2}$ 所界定的区域的体积.

12.3 计算 $\int_A (x^2+y^2)^2 dxdydz$, 其中 A 是由平面 $z = 2$ 和曲面 $x^2 + y^2 = 2z$ 所界定的区域 .

12.4 计算 $\int_A x$, 其中

$$A = \{\prec x, y, z \succ : x^2 + y^2 + z^2 \leqslant a^2, x \geqslant 0, y \geqslant 0, z \geqslant 0\} .$$

12.5 计算

$$\int_A \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right)^{1/2} ,$$

其中 A 是由椭球面

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

所界定的区域 .

设 ρ 是一个非负函数 (在下面讨论中称它为质量密度), 定义于 $\mathbb{E}^3$ 中一个区域 D 上. $\prec D, \rho \succ$ 的总质量被定义为

$$M = \int_D \rho(\boldsymbol{x}) d\boldsymbol{x} .$$

如果 $M \neq 0, \prec D, \rho \succ$ 的重心是点 $C = \prec C_1, C_2, C_3 \succ$, 其中

$$C_1 = \frac{1}{M}\int_D x_1 \rho(\boldsymbol{x}) d\boldsymbol{x} ,$$

$$C_2 = \frac{1}{M}\int_D x_2 \rho(\boldsymbol{x}) d\boldsymbol{x} ,$$

$$C_3 = \frac{1}{M}\int_D x_3\rho(\boldsymbol{x})d\boldsymbol{x} .$$

12.6 一个均匀立体 (即 ρ 为常数) 由 $x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0$ 和

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} + \frac{x_3^2}{c^2} \leqslant 1$$

给出. 求它的重心 .

12.7 单位立方体具有密度 $\rho(\boldsymbol{x}) = x_1 x_3$. 求其总质量和重心 .

12.8 求由曲面 $x^2+y^2+z^2=a^2$ 和 $x^2+y^2=ax$ 界定的均匀立体的质心 .

当然, 质心的概念可对任何维数的欧几里得空间中的区域定义. 因此平面上一个具有密度 ρ 的区域 D, 其质心是点 $\prec x_0, y_0 \succ$, 其中

$$x_0 = \frac{\int_D x\rho}{\int_D \rho} \quad 和 \quad y_0 = \frac{\int_D y\rho}{\int_D \rho} .$$

12.9 设 D 为 xz 平面上的区域, 整个位于半平面 $x>0$ 中. 设 A 为 $\mathbb{E}^3$ 中的立体, 由绕 z 轴旋转 D 得到. 证明 $\mu(A) = 2\pi d\mu(D)$, 其中 d 是从 D 的质心 (具均匀密度) 到 z 轴的距离.(利用柱坐标). 它以古尔丁(Guldin)定律而知名 .

考察重心的定义. 这时我们得到一个向量 (即 $\mathbb{E}^3$ 中一点) 当作答案, 这时我们对每个坐标在进行积分. 这提示了下面的定义: 设 V 为有限维向量空间, 并设 $e_1, \cdots, e_k$ 为 V 的一组基. 称从 $\mathbb{E}^n$ 到 V 的一个映射 f(f 是取值于 V 的 $\mathbb{E}^n$ 上的向量值函数) 为可度的是说, 如果当我们记 $f(\boldsymbol{x}) = \sum f_i(\boldsymbol{x})e_i$ 时, 函数 f_i 中每一个 (实值的) 都是可度的. 定义 f 在 D 上的积分为

$$\int_D f = \sum\left(\int_D f_i\right) e_i .$$

12.10 证明一个函数为可度的条件和它的积分值与基 $e_1, \cdots, e_k$ 的选取无关 .

设 ξ 是不在闭区域 D 中的一点, D 有一质量分布 ρ . 位于 ξ 的具单位质量的质点上的万有引力被定义为向量

$$\int_D \frac{\rho(\boldsymbol{x})(\boldsymbol{x}-\xi)}{\|\boldsymbol{x}-\xi\|^3} d\boldsymbol{x} .$$

(这里 $\boldsymbol{x}-\xi$ 是 $\mathbb{E}^3$ 上一个 $\mathbb{E}^3$ - 值的函数 .)

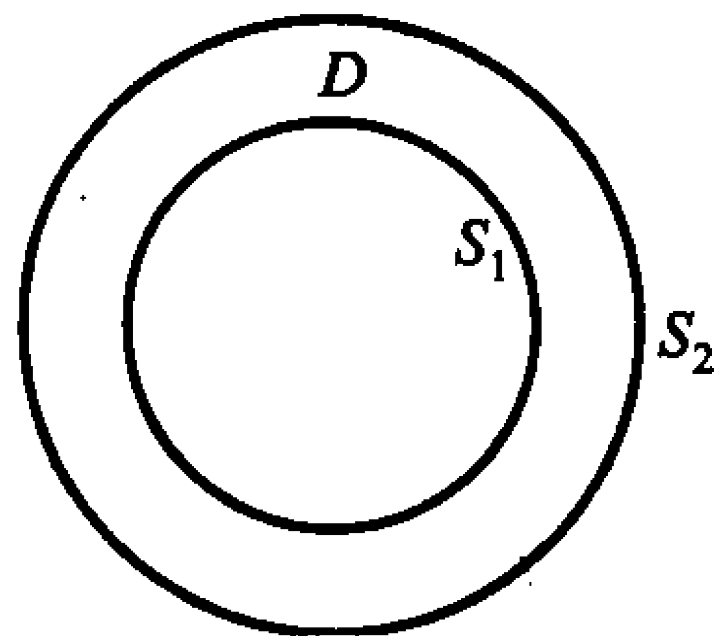


图 8.16

12.11 设 D 是个球壳, 由两个同心球面 S_1 和 S_2 所界定 (图 8.16), 中心在原点. 设 ρ 为 D 上的质量分布, 它只依赖于到中心的距离, 即 $\rho(\boldsymbol{x}) = f(\|\boldsymbol{x}\|)$. 证明万有引力在 S_1 内任何一点 ξ 均消失 .

12.12 $\prec D, \rho \succ$ 如习题 12.11 中的. 证明在 S_2 外一点上的万有引力与由一个位于原点并且质量为 D 的总质量的质点所产生的引力一样 .

8.13 绝对可积函数

迄今我们一直在处理具有紧支集的有界函数. 在实际应用上, 我们想要能够积分既非有界又不具紧支集的函数. 设 f 为定义于 $\mathbb{E}^n$ 上的函数, 设 M 为一非负实数, 并设 A 为 $\mathbb{E}^n$ 中一个 (有界的) 可度子集. 又设 f_A^M 为函数

$$f_A^M(x)=\begin{cases}0, & \text{如果 } x\notin A\,,\\ M, & \text{如果 } x\in A \text{ 且 } |f(x)|>M\,,\\ f(x), & \text{如果 } x\in A \text{ 且 } |f(x)|<M\,.\end{cases}$$

因此 f_A^M 是个具紧支集的有界函数. 它从 f 中在 A 外截 f 回到 0, 而当 $|f(x)|>M$ 则截 f 回到 M .

我们称一个函数 f 是绝对可积的说的是

(i) f_A^M 对所有 $M>0$ 和可度集 A 是可度函数 ;

(ii) 对任意 $\varepsilon>0$ 存在一个有界可度集 A_ε, 使得 $e_{A_\varepsilon}\cdot f$ 是有界的, 并且对所有 $M>0$ 和所有满足 $B\cap A_\varepsilon=\varnothing$ 的 B 有

$$\int|f_B^M|<\varepsilon\,.$$

容易验证两个绝对可积函数的和仍是绝对可积的. 因此绝对可积函数的集合构成了一个向量空间. 注意, 如果 f 满足条件 (i), 并且 $|f(x)|<|g(x)|$ 对所有 x 成立, 其中 g 是绝对可积的, 则 f 也绝对可积 .

设 f 是个绝对可积的函数. 给出任意 ε , 选取相应的 A_ε. 于是对任意 M_1 和 $M_2\geqslant \max_{x\in A_\varepsilon}|f(x)|$ 的数, 以及任意 $A_1\supset A_\varepsilon$ 和 $A_2\supset A_\varepsilon$ 的集合, 有

$$\left|\int f_{A_1}^{M_1}-\int f_{A_2}^{M_2}\right|\leqslant\int|f_{A_1}^{M_1}|+\int|f_{A_2}^{M_2}|<2\varepsilon\,.$$

如果让 $\varepsilon\to 0$ 并选取一个相应的 A_ε 族, 则上面的不等式表明 $\lim\int fe_{A_\varepsilon}$ 与 A_ε 的选取无关. 定义此极限为 $\int f$.

我们现在列出一个函数为绝对可积的一些简单自明的判别准则 . 我们要考虑两个不同的麻烦制造因素: 非有界性和没有紧支集 .

设 f 为有界函数, 使 f_A 对任意可度集 A 为可度. 假定对 $\|x\|$ 的大数值有 $|f(x)|\leqslant C\|x\|^{-k}$. 设 B_r 为中心在原点半径为 r 的球. 如果 r_1 足够大, 以致对 $\|x\|\geqslant r_1$ 时此不等式成立, 则对 $r_2\geqslant r_1$ 我们有

$$\int|fe_{B_{r_2}-B_{r_1}}|\leqslant C\int_{B_{r_2}-B_{r_1}}\|x\|^{-k}=C\Omega_n\int_{r_1}^{r_2}r^{n-1-k}\,,$$

其中 Ω_n 是某个依赖于 n 的常数 (事实上, 它是 $\mathbb{E}^n$ 中单位球面的 “曲面面积”). 如果 $k>n$, 最后面的那个积分变成为

$$\frac{C\Omega_n}{n-k}(r_2^{n-k}-r_1^{n-k}) ,$$

它 $\leqslant [C\Omega_n/(k-n)]r_1^{n-k}$, 当 $r_1\to\infty, k>n$ 时它趋向于 0. 因此我们可以断言:

设 f 是个有界函数使 fe_A 对任意可度集 A 为可度. 假设当 $\|x\|\to\infty$ 时 $|f(x)|\to 0$ 且使对某个 $k>n$ 有 $\|x\|^k|f(x)|$ 有界, 则 f 为绝对可积 .

现在我们来考虑当 f 具紧支集但无界的情形. 首先假定存在一个点 x_0 使得 f 在 x_0 的任一邻域的补集中有界. 进一步假定有某个常数 C 和 k, 使

$$|f(x)|\leqslant C^k\|x-x_0\|^{-k} .$$

于是, 如果 $|f(x)|>M$ 则 $\|x-x_0\|^{-k}>M/C^k$ 或者 $\|x-x_0\|\leqslant C/M^{1/k}$.

设 B_1 是半径为 $CM_1^{-1/k}$ 中心为 x_0 的球. 于是 $|f(x)|>M_1$ 表明 $x\in B_1$. 又, 对 $M_2>M_1$ 我们有

$$\int_{B_1}|f^{M_2}|<C^k\int_{B_1-B_2}\|x-x_0\|^k+M_2\mu(B_2) ,$$

其中 B_2 是半径为 $CM_2^{-1/k}$ 中心在 x_0 的球. 因此

$$\int_{B_1}|f^{M_2}|\leqslant C^k\Omega_n\int_{CM_2^{-1/k}}^{CM_1^{-1/k}} r^{n-1-k}dr+M_2C^{n-1}V_nM_2^{-n/k} ,$$

其中的 Ω_n 和 V_n 只依赖于 n. 如果 $k<n$, 右端的积分成为

$$\frac{C^{n-k}}{n-k}(M_1^{(k-n)/k}-M_2^{(k-n)/k})<\frac{C^{n-k}}{n-k}M_1^{(k-n)/k} .$$

因此

$$\int_{B_1}|f^{M_2}|<\mathrm{const}\cdot(M_1^{(k-n)/k}+M_2^{(k-n)/k}) ,$$

在选取 M_1 充分大时可以使它充分小 .

从而若 f 具紧支集并对所有 M 有 f^M 可度, 以及当 $k<n$ 时 $|f(x)|<C\|x-x_0\|^{-k}$, 那么 f 则是绝对可积的 .

更一般地, 设 S 为 $\mathbb{E}^n$ 的一个 l 维子空间的有界子集. 设 $d(x)$ 表示从 x 到 S 的距离. 设 f 是对所有 M 有 f^M 可度的具紧支集的函数. 如果 $|f(x)|<Cd(x)^{-k}$, 其中 $k<n-l$, 则 f 是绝对可积的. 其证明与上面给出的相似, 留给读者 .

设 $\{f_k\}$ 为一个绝对可积函数的序列. 如果序列 $f_k(x) \to f(x)$, 那么在什么条件下才有 $\int f_n \to \int f$? 即使是此序列一致收敛也不能保证积分收敛. 例如, 如果 $f_k = (1/k^n)e_{\square_0^k}$, 则 $|f_k(x)| \leqslant \frac{1}{k^n}$, 故 f_k 一致地趋向零. 另一方面, 对所有 k 有 $\int f_k = 1$.

我们说函数集 $\{f_k\}$一致绝对可积是指, 如果对任意 $\varepsilon > 0$, 存在一个 A_ε, 它可以以独立于k 的方式选取, 使得只要 $B \cap A_\varepsilon = \varnothing$, 对所有 M 有

$$\int_B |f_k^M| < \varepsilon .$$

我们经常用来验证 $\{f_k\}$ 为一致绝对可积的办法是证明存在一个绝对可积函数 g, 使得 $|f_k(x)| \leqslant |g(x)|$ 对所有的 k 和 x 都成立 .

设 $\{f_k\}$ 为函数的一个一致绝对可积序列. 假设 $f_k \to f$ 是一致的. 另外再假设 f 是绝对可积的. 于是有 $\int f_k \to \int f$. 事实上, 对任意 $\delta > 0$ 我们可以找到 k_0 使得对所有 $k > k_0$ 和所有 x 有 $|f_k(x) - f(x)| < \delta$. 我们也可找到 A_ε 和 M_ε, 使得

$$\begin{aligned}\left|\int f_k - \int f\right| &\leqslant \left|(f_k)_{A_\varepsilon}^M - \int f_k\right| + \left|f_{A_\varepsilon}^M - \int f\right| + \left|(f_k)_{A_\varepsilon}^M - \int f_{A_\varepsilon}^M\right| \\ &\leqslant \varepsilon + \varepsilon + \delta\mu(A_\varepsilon) ,\end{aligned}$$

这是可以做成任意小的, 只要先选取小的 ε(从而给出了 A_ε) 然后选取小的 δ(意味着选取了大的 k_0) .

上述这些想法的主要应用是在计算累次积分问题中和在积分号下取微分的问题中 .

命题 13.1 设 f 是 $\mathbb{R}^k \times \mathbb{R}^l$ 上的函数. 假定函数集 $\{f(x, \cdot)\}$ 一致绝对可积, 其中 x 被限制于一个有界可度集 $K \subset \mathbb{R}^k$ 中. 于是函数 $e_{K\times\mathbb{R}^l} \cdot f$ 是绝对可积的, 并且

$$\int_{K\times\mathbb{R}^l} f = \int_K \int_{\mathbb{R}^l} f(x,y)dydx = \int_{\mathbb{R}^l} \int_K f(x,y)dxdy .$$

证明 由假定条件, 对任意 $\varepsilon > 0$, 可以找到 M 和 $A_\varepsilon \subset \mathbb{R}^l$ 使得, 如果 $A \cap A_\varepsilon = \varnothing$

$$\int_{\mathbb{R}^l} |f_A^M(x, \cdot)| < \varepsilon . \tag{13.1}$$

现在对任意 $\mathbb{R}^n$ 中的集合 B 有

$$\int e_{K\times\mathbb{R}^l}|f_B^M| = \int e_K(x)\int_{\mathbb{R}^l}|f_B^M(x,\cdot)|$$
$$\leqslant \mu(K)\varepsilon$$

其中 $B\cap K\times A_\varepsilon=\varnothing$.

这表明 $e_{K\times\mathbb{R}^l}f$ 在 $\mathbb{R}^n$ 上是绝对可积的. 现选取充分大的 $\square=\square_1\times\square_2$ 和一个 M 使得

$$\left|\int_{K\times\mathbb{R}^l}f-\int_{K\times\mathbb{R}^l}f_\square^M\right|<\varepsilon ,$$

也使得对所有 $x\in K$ 有

$$\left|\int f(x,\cdot)-\int f_\square^M(x,\cdot)\right|<\varepsilon .$$

于是我们有了

$$\left|\int_{K\times\mathbb{R}^l}f-\int_{K\times\mathbb{R}^l}f_\square^M\right|<\varepsilon$$

和

$$\int_{K\times\mathbb{R}^l}f_\square^M=\int_K\int_{\mathbb{R}^l}f_B^M(x,y)dydx .$$

因此

$$\left|\int_{K\times\mathbb{R}^l}f-\int_K\int_{\mathbb{R}^l}f(x,y)dydx\right|<\varepsilon+\mu(K)\varepsilon ,$$

故

$$\int_{K\times\mathbb{R}^l}f=\int_K\int_{\mathbb{R}^l}f(x,y)dydx .$$

最后, 等式 (13.1) 表明函数 $F(y)=\int_K f(\cdot,y)$ 是绝对可积的. 事实上, 用 (13.1) 式中同一个 A 和 M, 我们得到

$$\int|F_B^{M_1}|\leqslant\mu(K)\varepsilon .$$

因而我们得到了

$$\int_{K\times\mathbb{R}^l}f=\int_K\int_{\mathbb{R}^l}f(x,y)dydx=\int_{\mathbb{R}^l}\int_K f(x,y)dxdy . \qquad \square$$

以相同论证得到下面的一个推广 .

命题 13.2 设 f 在 $\mathbb{R}^n$ 上绝对可积并使得函数 $f(x,\cdot)$ 对每个 $x\in\mathbb{R}^k$ 为一致绝对可积，则

$$\int f = \iint f(x,y)dydx\ .$$

我们现在把注意力转向积分号下求微分的问题．

命题 13.3 设 $(t,x)\mapsto F(t,x)$ 为 $I\times\mathbb{R}^n$ 上的函数，其中 $I=[a,b]\subset\mathbb{R}$. 假设

(i) F 和 $\partial F/\partial t$ 在 $I\times\mathbb{R}^n$ 上为连续函数；

(ii) $(\partial F/\partial t)(t,\cdot)$ 是函数的一个一致绝对可积族；

(iii) $F(t,\cdot)$ 对所有 $t\in I$ 是绝对可积的．

设 $f(t)=\int F(t,\cdot)$. 于是 f 是 t 的可微函数且

$$f'(t)=\int_{\mathbb{R}^n}(\partial F/\partial t)(t,\cdot)$$

证明 设 $G(t)=\int_{\mathbb{R}^n}(\partial F/\partial t)(t,\cdot)$. 则 $G(t)$ 为连续; 从而对在积分号下的一族绝对可积函数我们可使其过渡到极限. 进一步由命题 13.1 还有

$$\int_a^t G(s)ds=\int_{\mathbb{R}^n}\int_a^t(\partial F/\partial t)(s,\cdot)ds\ .$$

因此,

$$\begin{aligned}\int_a^t G(s) &= \int_{\mathbb{R}^n}(F(t,\cdot)-F(a,\cdot))=\int_{\mathbb{R}^n}F(t,\cdot)-\int_{\mathbb{R}^n}F(a,\cdot)\\ &= f(t)-f(a)\ .\end{aligned}$$

对此方程作关于 t 的微分便得出所要的结果． □

最后, 我们叙述对绝对可积函数的换元公式．

命题 13.4 设 $\varphi:U\to V$ 为可微的一一映射并且有可微逆，其中 U 和 V 为 $\mathbb{R}^n$ 中两个开集. 设 f 为定义于 V 上的一个绝对可积函数. 则 $(f\circ\varphi)|\det J_\varphi|$ 是 U 上的绝对可积函数，并且

$$\int_V f=\int_U(f\circ\varphi)|\det J_\varphi|\ .$$

证明 要证明 $(f\circ\varphi)|\det J_\varphi|$ 为绝对可积, 我们设 $\varepsilon>0$ 并选取 $A_\varepsilon\subset V$ 使得 (ii) 成立. 于是 $\overline{A}_\varepsilon$ 为紧. 因而 $\varphi^{-1}(\overline{A}_\varepsilon)$ 也紧. 特别, $\varphi^{-1}(\overline{A}_\varepsilon)$ 是个有界可度集且

$|\det J_\varphi|$ 在它上面有界. 如果 $B\cap\varphi^{-1}(\overline{A}_\varepsilon)=\varnothing$, 其中 $B\subset U$ 为有界和可度的, 则

$$\int_B |\det J_\varphi|^M|(f\circ\varphi)^M| \leqslant \int_{\varphi(B)} |f^M| < \varepsilon\,.$$

这表明 $(f\circ\varphi)|\det J_\varphi|$ 为绝对可积. 命题的其余部分从

$$\int_{A_\varepsilon} f = \int_{\varphi^{-1}(A_\varepsilon)} (f\circ\varphi)|\det J_\varphi|$$

中让 $\varepsilon\to 0$ 得到. □

习题

13.1 计算积分 $\int_{-\infty}^{\infty} e^{-x^2}dx$. [提示: 计算其平方.]

13.2 计算积分 $\int_0^{\infty} e^{-x^2}x^{2k}dx$.

13.3 计算在一个奇维数空间中单位球的体积.[提示: 看出对"极坐标"的雅可比行列式具有 $r^{n-1}\times f$ 的形式, 其中 f 是个"角变量"的函数. 因此单位球的体积具有形式 $C\int_0^1 r^{n-1}dr$, 其中 C 由 f 在"角变量"上的积分所决定. 由计算 $(\int_{-\infty}^{\infty} e^{-x^2}dx)^n$ 来得出 C 的值.]

8.14 问题汇编: 傅里叶变换

设 $\alpha=\prec\alpha_1,\cdots,\alpha_n\succ$ 为一个 n 联组, 其分量为非负整数. 用 D^α 表示微分算子

$$D^\alpha=\frac{\partial^{\alpha_1+\cdots+\alpha_n}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}\,.$$

设 $|\alpha|=\alpha_1+\cdots+\alpha_n$. 又设 $Q(x,D)=\sum_{|\alpha|\leqslant k}a_\alpha(x)D^\alpha$ 为微分算子, 其中每个 a_α 是 x 的多项式. 因此如果 f 是 $\mathbb{R}^n$ 上的一个 C^k 函数, 我们有

$$(Qf)(x)=\sum_{|\alpha|<k}a_\alpha(x)D^\alpha f(x)\,.$$

对任何在 $\mathbb{R}^n$ 上的 C^∞ 函数, 我们令

$$\|f\|_Q=\sup_{x\in\mathbb{R}^n}|Qf(x)|\,.$$

我们以 $\mathcal{S}$ 表示所有 $f \in C^\infty$, 它们对所有 Q 满足

$$\|f\|_Q < \infty \tag{14.1}$$

所组成的空间. 要明白它的意思, 让我们考虑那些 $k = 0$ 的 Q. 于是 (14.1) 式说, 对任意多项式 $a(\cdot)$, 函数 $a \cdot f$ 是有界的. 换句话说, f 在无穷远处比任何多项式的逆消失得更快; 即对所有 p

$$\lim_{\|x\| \to \infty} \|x\|^p f(x) = 0 .$$

说 (14.1) 式成立意味着上面的等式对 f 的任何导数也成立 .

如果 f 是个具紧支集的 C^∞ - 函数, 则 (14.1) 式显然成立, 故 $f \in \mathcal{S}$. 一个更富意义的例子是由函数 n 给出的, 这里

$$n(x) = e^{-\|x\|^2} .$$

因为 $\lim_{r\to\infty} r^p e^{-r^2} = 0$ 对所有的 p 成立, 从而推出 $\lim_{\|x\|\to\infty} a(x)n(x) = 0$. 另一方面, 容易看出 (由归纳法)$D^\alpha n(x) = P_\alpha(x)n(x)$, 其中 P_α 为某个多项式. 因此 $Qn(x) = P_Q(x)n(x)$, 其中 P_Q 为多项式. 故 $n \in \mathcal{S}$.

容易看出空间 $\mathcal{S}$ 是个向量空间. 我们将在此空间上引进收敛性的概念, 即称 $f_n \to f$ 是说对每个固定的 Q, 有

$$\|f_n - f\|_Q \to 0 .$$

(注意, 空间 $\mathcal{S}$ 不是巴拿赫空间, 因为在其中的收敛性依赖于无穷多个不同的范数 .)

习题

14.1 设 φ 为一个 C^∞ 函数, 它在无穷处增长缓慢. 就是说, 假设对每个 α 有一个多项式 P_α 使得对所有 x 有

$$|D^\alpha \varphi(x)| < P_\alpha(x) .$$

说明如果 $f \in \mathcal{S}$, 则 $\varphi f \in \mathcal{S}$. 另外由 $f \mapsto \varphi f$ 定义的 $\mathcal{S}$ 到自己的映射是连续的, 即如果 $f_n \to f$ 则 $\varphi f_n \to \varphi f$.

对 $x = \prec x^1, \cdots, x^n \succ \in \mathbb{R}^n, \xi = \prec \xi^1, \cdots, \xi^n \succ \in \mathbb{R}^{n*}$, 我们用

$$\prec x, \xi \succ = x^1\xi^1 + \cdots + x^n\xi^n$$

代表 ξ 在 x 的值. 又对于任意 $\alpha = \prec \alpha^1, \cdots, \alpha^n \succ$ 及任意 $x \in \mathbb{R}^n$, 我们令

$$x^\alpha = (x^1)^{\alpha_1} \cdots (x^n)^{\alpha_n} ,$$

相似地有 $\xi^\alpha = (\xi^1)^{\alpha_1} \cdots (\xi^n)^{\alpha^n}$, 等等.

对任意 $f \in \mathcal{S}$, 定义它的傅里叶变换 $\hat{f}$ 为 $\mathbb{R}^{n*}$ 上的函数,

$$\hat{f}(\xi) = \int e^{-i\prec x,\xi \succ} f(x)dx .$$

我们看到有

$$\hat{f}(0) = \int f \quad 和 \quad |\hat{f}(\xi)| \leqslant \int |f| .$$

14.2 证明 $\hat{f}$ 具有关于 ξ 的任意阶导数, 并且

$$D_\xi^\alpha \hat{f}(\xi) = (-i)^{|\alpha|} \int e^{-i\prec x,\xi \succ} x^\alpha f(x)dx ;$$

换句话说,

$$D_\xi^\alpha \hat{f}(\xi) = \hat{g}(\xi) ,$$

其中 $g(x) = (-i)^{|\alpha|} x^\alpha f(x)$.

14.3 证明

$$\widehat{\frac{\partial f}{\partial x^j}}(\xi) = i\xi^j \hat{f}(\xi) .$$

[提示: 将此积分写为累次积分并用关于第 j 个变量的分部积分法 .]

14.4 得出下面结论: 映射 $f \mapsto \hat{f}$ 将 $\mathcal{S}(\mathbb{R}^n)$ 映到 $\mathcal{S}(\mathbb{R}^{n*})$, 并且如果在 $\mathcal{S}$ 中 $f_n \to 0$ 则在 $\mathcal{S}(\mathbb{R}^{n*})$ 中 $\hat{f}_n \to 0$.

14.5 证明 $\widehat{T\omega f}(\xi) = e^{-i\prec \omega,\xi \succ} \hat{f}(\xi)$ 对任意 $\omega \in \mathbb{R}^n$ 成立. 回想这里的 $T_\omega f(x) = f(x-\omega)$.

14.6 对任意 $f \in \mathcal{S}$, 定义 $\tilde{f}$ 为

$$\tilde{f}(x) = \overline{f(-x)} ,$$

其中 — 表示复共轭. 证明

$$\hat{\tilde{f}}(\xi) = \overline{\hat{f}(\xi)} .$$

14.7 设 $n = 1$, 并设 f 为 x 的偶的实函数. 证明

$$\hat{f}(\xi) = \int \cos \prec x, \xi \succ f(x)dx .$$

14.8 设 $n(x) = e^{-(1/2)x^2}$, 其中 $x \in \mathbb{R}^1$. 证明

$$\frac{d\hat{n}}{d\xi}(\xi) = -\xi \hat{n}(\xi) ,$$

并有结论

$$\log \hat{n}(\xi) = -\frac{1}{2}\xi^2 + 常数$$

从而

$$\hat{n}(\xi) = 常数 \times e^{-(1/2)\xi^2} .$$

由置 $\xi = 0$ 并用习题 13.1, 计算出这个常数为 $\sqrt{2\pi}$. 因此 $\hat{n}(\xi) = \sqrt{2\pi} e^{-(1/2)\xi^2}$.

14.9 证明极限 $\lim_{\epsilon \to 0} \int_{\epsilon}^{1/\epsilon} (\sin\ x)/x$ 存在. 记此极限为 d. 证明对任意 $R > 0$, $\lim_{\epsilon \to 0} \int_{\epsilon}^{1/\epsilon} (\sin\ Rx)/x dx = d$.

如果 $f \in \mathcal{S}$, 我们已知有 $\hat{f} \in \mathcal{S}(\mathbb{R}^{n*})$. 因此我们可以考虑函数

$$\int e^{i \prec y, \xi \succ} \hat{f}(\xi) d\xi .$$

下面几个习题的目的是要证明

$$f(y) = \frac{1}{(2\pi)^n} \int e^{i \prec y, \xi \succ} \hat{f}(\xi) d\xi . \tag{14.2}$$

我们首先注意到, 由于所涉及到的所有积分是绝对收敛的, 故只须证明

$$f(y) = \lim_{R_1 \to \infty} \cdots \lim_{R_n \to \infty} \frac{1}{(2\pi)^n} \cdot \int_{-R_n}^{R_n} \cdots \int_{-R_1}^{R_1} \hat{f}(\xi^1, \cdots, \xi^n) e^{i(y'\xi' + \cdots + y^n \xi^n)} d\xi^1 \cdots d\xi^n .$$

将 $\hat{f}$ 的定义代入此公式并交换关于 x 和 ξ 进行积分的顺序, 我们便得到

$$\lim_{R_1 \to \infty} \cdots \lim_{R_n \to \infty} \left(\frac{1}{2\pi}\right)^n \iint_{-R_n}^{R_n} \cdots \int_{-R_1}^{R_1} f(x^1, \cdots, x^n) e^{i[(y^1 - x^1)\xi^1 + \cdots + (y^n - x^n)\xi^n]} d\xi^1 \cdots d\xi^n dx .$$

因此只需一次计算一个变量的极限就行了 (假定了此收敛性是一致的, 这可由证明中弄清楚). 我们于是把此问题化为单变量函数情形. 我们必须证明, 如果 $f \in \mathcal{S}(\mathbb{R}^1)$, 则

$$f(y) = \lim_{R \to \infty} \frac{1}{2\pi} \iint_{-R}^{R} f(x) e^{i(y-x)\xi} d\xi dx .$$

我们将首先证明

$$f(y) = \lim_{R \to \infty} \frac{1}{4d} \iint_{-R}^{R} f(x) e^{i(y-x)\xi} d\xi dx .$$

其中 d 已在习题 14.9 中给出 .

14.10 证明上面最后一个积分可以写为

$$\frac{1}{2d}\int_{-\infty}^{\infty} f(x)\frac{\sin R(y-x)}{y-x}dx = \frac{1}{d}\int_{0}^{\infty} \frac{f(y-x)+f(y+u)}{2}\sin\frac{Ru}{u}du\ .$$

14.11 设

$$g(u) = \frac{f(y-u)+f(y+u)}{2} - f(y)\ .$$

证明 $g(0)=0$ 并得到对 $0 \leqslant x \leqslant 1$ 有 $g(x) = xh(x)$, 其中 $h \in C^1$. 用分部积分法证明

$$\left|\int_1^{\varepsilon} g(u)\sin\frac{Ru}{u} + \int_{1/\varepsilon}^{1} g(u)\sin\frac{Ru}{u}\right| < \text{const}\cdot\frac{1}{R}\ .$$

最后有

$$\lim_{R\to\infty}\int_0^{\infty}\frac{f(y-u)+f(y+u)}{2}\sin\frac{Ru}{u}du = f(y)\ .$$

这便证明了

$$f(y) = \frac{1}{4d}\int e^{iy\xi}f(\xi)d\xi\ .$$

14.12 用习题 14.8 得出 $d=\pi/2$.

设 $f_1 \in \mathcal{S}, f_2 \in \mathcal{S}$. 定义函数 $f_1 \star f_2$ 为

$$f_1 \star f_2(x) = \int f_1(x-y)f_2(y)dy\ .$$

由于右边的被积项显然对每个固定的 x 值收敛, 故此表达式有明确的意义. 我们还能使其更准确. 因为 $f_i \in \mathcal{S}$, 对任一整数 p, 我们可以找到一个 K_p 使得

$$|f_i(y)| < \frac{K_p}{1+\|y\|^p}\ ,$$

故

$$\int_{\|y\|>R}|f_2(y)| < \frac{L_pR^n}{1+R^p}\ .$$

于是

$$\int(1+\|x\|^q)f_1(x-y)f_2(y)dy = \int_{\|y\|<(1/2)\|x\|}(1+\|x\|^q)f_1(x-y)f_2(y)dy + \int_{\|y\|>(1/2)\|x\|}(1+\|x\|^q)f_1(x-y)f_2(y)dy\ .$$

第一个积分最多为

$$C_n\left(\frac{1}{2}\|x\|\right)^n(1+\|x\|)^q\max_z|f_2(z)|\frac{K_p}{1+(\frac{1}{2}\|x\|)^p}\ ,$$

而第二个最多为

$$(1+\|x\|^q)\max_u|f_1(u)|\frac{L_p\left(\frac{1}{2}\|x\|\right)^n}{1+\left(\frac{1}{2}\|x\|\right)^p}\ .$$

选取 $p>q+n$, 我们看出两项都趋向零. 因此

$$\lim_{\|x\|\to\infty}(1+\|x\|^q)f_1\star f_2(x)=0\ .$$

14.13 证明

$$\frac{\partial}{\partial x^i}(f_1\star f_2)=\left(\frac{\partial f_1}{\partial x^i}\right)\star f_2=f_1\star\left(\frac{\partial f_z}{\partial x^i}\right)\ .$$

从而得出 $f_1\star f_2\in\mathcal{S}$.

14.14 证明如果 φ 是 $\mathbb{R}^n$ 上任一有界连续函数, 则

$$\iint\varphi(x+y)f_1(x)f_2(y)dxdy=\int\varphi(u)(f_1\star f_2)(u)du\ .$$

14.15 得出结论

$$\widehat{f_1\star f_2}(\xi)=\hat{f}_1(\xi)\hat{f}_2(\xi)\ .$$

14.16 证明

$$f\star\tilde{f}(y)=\left(\frac{1}{2\pi}\right)^n\int|\hat{f}(\xi)|^2e^{i\prec y,\xi\succ}d\xi\ .$$

14.17 对任意 $f\in\mathcal{S}$ 得出结论说

$$\int|f|^2=\left(\frac{1}{2\pi}\right)^n\int|\hat{f}|^2\ . \tag{14.3}$$

[提示: 在习题 14.16 中令 $y=0$.]

下面的习题运用了傅里叶变换讲述那些在研究偏微分方程中有用的东西. 在最后一章的末尾我们要利用它们. 读者可以选择一直推延到那时再学习这些问题 .

在空间 $\mathcal{S}$ 上定义范数 $\|\ \|_s$ 为

$$\|f\|_s^2=(2\pi)^{-n}\int(1+\|\xi\|^2)^s|\hat{f}(\xi)|^2d\xi\ ,$$

而内积 $(f,g)_s$ 为

$$(f,g)_s=\int(1+\|\xi\|^2)^s\hat{f}(\xi)\widehat{g}(\xi)d\xi\ .$$

14.18 设 $s=R$ 为非负整数. 证明

$$\|f\|_s^2=\sum_{|\alpha|\leqslant R}\frac{R}{\alpha!(R-|\alpha|)!}\int|D^\alpha f(x)|^2dx\ ,$$

其中 $\alpha!=\alpha_1!\cdots\alpha_n!$. [利用多项式展开定理, 重复应用习题 14.3 和等式 (14.3)]

我们因而看到 $\|f\|_R$ 以平方的积分范数度量了 f 和它直到 R 阶导数的大小. 它有助于把 $\| \ \|_s$ 想成为大小概念的一种推广, 它推广到 s 可以为任意实数的情形.

注意如果 $s<t$, 我们有

$$\|f\|_s \leqslant \|f\|_t .$$

对于任意实的 s, 定义算子 K^s 为

$$\widehat{K^s f}(\xi) = (1+\|\xi\|^2)^s \hat{f}(\xi) .$$

14.19 证明由

$$Kf = f - \sum \frac{\partial^2 f}{\partial x_i^2}$$

给出的算子 $K = K^1$.

14.20 证明对任意实数 s 和 t.

$$\|K^s f\|_t = \|f\|_{t+2s}$$

以及

$$(K^s f, g)_t = (f, K^s g)_t = (f, g)_{s+t} .$$

14.21 证明 $K^{s+t} = K^s \circ K^t$, 故特别有对所有 s, K^s 为可逆.

我们现在定义空间 H_s 为 $\mathcal{S}$ 在范数 $\| \ \|_s$ 下的完备化. 空间 H_s 是具内积 $(,)_s$ 的希尔伯特空间. 我们可以把 H_s 的元素想成是"具直到 s 阶的广义导数的广义函数". 由构造, 空间 $\mathcal{S}$ 是 H_s 在范数 $\| \ \|_s$ 下的一个稠子空间. 我们注意到, 习题 14.20 表明算子 K^s 可以扩张为 H_t 到 H_{t-2s} 的一个等距映射. 我们也把此扩张的映射记为 K^s. 由习题 14.21,

$$K^{-s} : H_{t-2s} \ \to \ H_t$$

是 K^s 的逆, 故 K^s 是 H_t 到 H_{t-2s} 上的一个保范数同构.

14.22 设 $\mu \in H_s, v \in H_{-s}$. 证明

$$|(u,v)_0| \leqslant \|\mu\|_s \|v\|_{-s} .$$

因此我们可以把 $\prec \mu, v \succ \mapsto (u,v)_0$ 扩张到 $H_s \times H_{-s}$ 上的一个函数, 它对 u 线性而对 v 为反线性 [即 $(u, av_1 + bv_2)_0 = \overline{a}(u,v_1) + \overline{b}(u,v_2)$] 并且满足上面的不等式. 从而任意的 $v \in H_{-s}$ 定义了一个有界线性函数 l 于 H_s 上, 它由 $l(u) = (u,v)_0$ 给出.

14.23 反之, 设 l 是 H_s 上一个有界线性函数. 证明存在一个 $v \in H_{-s}$, 满足 $l(u) = (u,v)$. 对所有 $u \in H_s$ 成立. [提示: 考虑线性形式 $v = K^s w$, 其中 w 为 H_s 中一个适当的元素; 应用第五章的定理 2.4 .]

14.24 证明

$$\|v\|_{-s}=\sup_{\substack{u\in H_s\\u\neq 0}}\frac{|(u,v)_0|}{\|u\|_s}.$$

(习题 14.22 给出了一个不等式. 如果 $v\neq 0$, 取 $u=K^{-s}v$ 能得到

$$\|v\|_{-s}\|u\|_s=(K^{-s/2}v,K^{-s/2}v)=\|v\|_s^2,$$

这样做是为了得出一个等式.)

14.25 设 $2s>n$(其中我们的函数定义于 $\mathbb{R}^n$ 上). 证明对任意 $f\in\mathcal{S}$ 我们有

$$\sup_{\mathbb{R}^n}|f(x)|\leqslant\|f\|_s\left[\int_{\mathbb{R}^n}(1+\|\xi\|^2)^{-s}\right]^{1/2}d\xi\quad(\text{索伯列夫不等式}).$$

(利用等式 (14.2), 施瓦兹不等式, 以及不等式右端的积分是绝对收敛这个事实 .)

索伯列夫不等式证明了 $\mathcal{S}$ 到 $C(\mathbb{R}^n)$ 的内射扩张为 H_s 到 $C(\mathbb{R}^n)$ 内的连续内射, 其中在 $C(\mathbb{R}^n)$ 上给出的是一致范数. 因此我们可以把 H_s 的元素在 $s>n/2$ 时看成实际上是 $\mathbb{R}^n$ 上的函数 .

对 $|\alpha|$ 由归纳法我们便可断言, 对 $s>n/2$, 任意 $f\in H_{|\alpha|+s}$ 具有 $|\alpha|$ 阶连续导数, 并且

$$\sup_{\mathbb{R}^n}|D^\alpha f(x)|\leqslant C\alpha\|f\|_{|\alpha|+s}.\tag{14.4}$$

14.26 设 Ω 为 $\mathbb{R}^n$ 的一个有界开子集. 设 $\varphi\in\mathcal{S}$ 满足 $\operatorname{supp}\varphi\subset\Omega$. 证明对所有 ξ, 有

$$|\hat{\varphi}(\xi)|\leqslant\mu(\Omega)^{1/2}\|\varphi\|_0$$

14.27 设 $d^i=\operatorname{lub}_{x\in\Omega}|x^i|$, 并设 $d^\alpha=(d^1)^{\alpha_1}\cdots(d^n)^{\alpha_n}$. 证明

$$|D^\alpha\hat{\varphi}(\xi)|\leqslant d^\alpha\mu(\Omega)^{1/2}\|\varphi\|_0.$$

14.28 证明

$$|\xi^\beta\hat{\varphi}(\xi)|\leqslant\mu(\Omega)^{1/2}\|D^\beta\varphi\|_0,$$

并最后得到

$$|(1+\|\xi\|^2)^k\hat{\varphi}^2(\xi)|\leqslant\mu(\Omega)\|\varphi\|_k^2.$$

14.29 更一般地, 设 ψ 为 $\mathcal{S}$ 中函数, 满足对所有 $x\in\Omega$ 有 $\psi(x)=1$, 又设 $\varphi\in\mathcal{S}$ 满足 $\operatorname{supp}\varphi\subset\Omega$. 证明

$$|\hat{\varphi}(\xi)|=|(\varphi,\psi_\xi)_0|\leqslant\|\varphi\|_s\|\psi_\xi\|_{-s},$$

其中 $\psi_\xi(x)=\psi(x)e^{-i\prec x,\xi\succ}$, 并且

$$|D_\xi^\alpha\hat{\varphi}(\xi)|\leqslant\|\varphi\|_s\|\psi_\xi^\alpha\|_{-s},$$

其中 $\psi_\xi^\alpha(x) = x^\alpha \psi(x) e^{-i\prec x,\xi\succ}$.

我们以 H_s^Ω 表示 $\mathcal{S}$ 中那些支集在 Ω 中的函数空间在 $\|\ \|_s$ 下的完备化. 按照习题 14.29, 任意 $\varphi \in H_s^\Omega$ 定义了一个 ξ 的函数 $\hat{\varphi}$, 它是可微的并满足

$$|D_\xi^\alpha \hat{\varphi}(\xi)| \leqslant \|\varphi\|_s \|\psi_\xi^\alpha(x)\|_{-s} ,$$

其中 $\|\psi_\xi^\alpha(x)\|_{-s}$ 只依赖于 Ω, α, ξ 和 $-s$ 而与 φ 无关. 进一步还有 $\|\varphi\|_s^2 = \int (1+\|\xi\|^2)^s |\hat{\varphi}(\xi)|^2 d\xi$.

14.30 设 $s < t$. 于是内射 $H_t \to H_s$ 是个紧映射. 就是说, 如果 $\{\varphi_i\}$ 是 H_t^Ω 中元素的序列使得对所有 i 有 $\|\varphi_i\|_t \leqslant 1$, 则我们可以选择一个子序列 $\{\varphi_{i_j}\}$, 它在 $\|\ \|_s$ 下收敛. [提示:由习题 14.29, 函数序列 $\{\varphi_i(\xi)\}$ 在 $\{\xi : \|\xi\| \leqslant r\}$, 固定的 r 上是有界和等度连续的. 我们因而可以选取一个子序列它一致收敛, 并由此有一个子子序列在 $\{\xi : \|\xi\| < r\}$ 对所有的 r 收敛 (一致性可能依赖于 r). 于是, 如果 $\{\varphi_{i_j}\}$ 是这个子子序列, 便有

$$\begin{aligned}
\|\varphi_{i_j} - \varphi_{i_k}\|_s^2 &= \int (1+\|\xi\|^2)^s |\varphi_{i_j}(\xi) - \varphi_{i_k}(\xi)|^2 d\xi \\
&= \int_{\|\xi\| \leqslant r} (1+\|\xi\|^2)^s |\varphi_{i_j}(\xi) - \varphi_{i_k}(\xi)|^2 d\xi \\
&\quad + \int_{\|\xi\| > r} (1+\|\xi\|^2)^s |\varphi_{i_j}(\xi) - \varphi_{i_k}(\xi)|^2 d\xi \\
&\leqslant \int_{\|\xi\| \leqslant r} (1+\|\xi\|^2)^s |\varphi_{i_j}(\xi) - \varphi_{i_k}(\xi)|^2 d\xi \\
&\quad + (1+\|\xi\|^2)^{s-t} \{\|\varphi_{i_j}\|_t^2 + \|\varphi_{i_k}\|_t^2\}].
\end{aligned}$$

第九章 微分流形

迄今我们对微积分的研讨一直在致力于对定义在一个向量空间 (或它的子集上) 函数的性质及在这些函数上的算子的研究. 用到的主要思想是对在每点可能是非线性的函数使用线性函数去逼近. 在这一章里我们要推广空间的概念使其包含那些不能以任何自然的方式被看作向量空间中开子集的空间. 我们要使用的工具是在每点用线性空间去"逼近"这类空间.

假定我们有意于研究定义在 $\mathbb{E}^3$ 中的单位球面 (表面的) 上的函数. 这个球面是个二维的数学对象, 它的意思是说对在此球面的每点的一个邻域我们可以用两个坐标以双向连续的方式来描述. 另一方面, 我们却不能以双向连续的一对一方式把一个球面映成平面的一个开子集 (因为球面是紧的而 $\mathbb{E}^2$ 中开子集则不是). 因而球面的一块可以用 $\mathbb{E}^2$ 中开子集来描述但整个球则不行. 所以, 如果我们要同时在整个球面上做微积分, 我们就必须引进更加广泛的空间类并研究在它们上面的函数.

即便一个空间可以被看成某个向量空间的子集, 也可想见它不能以一种典型的方式这样去看. 那么, 一个 (均匀的理想) 气体的平衡态可以被三个参数: 温度, 压力和体积中任意两个来阐述. 没有任何理由要选择这两个而不选择第三个. 从一组参数转移到另一组的变换是由一个一对一的双向可微映射给出. 因此此气体状态的任何函数在一种参数选择下是可微函数也在另一种选择下可微. 所以谈论关于气体状态的可微函数是有明确意义的. 然而, 一个函数对一种选择

的参数是线性的则不必对另一种选择的参数为线性. 那么谈论这个气体状态的线性函数的确没有任何意义. 在这种情况下, 我们想知道的是在此空间中函数的什么性质和什么运算是有意义的, 而不是那些我们给出此空间的那种人为制造的描述.

最后, 甚至有时为方便解决某些特定问题而在向量空间中引进了"非线性坐标": 例如, 在第八章习题 11.3 和 11.4 中的极坐标. 因此我们想要知道当我们改变坐标时各种数学对象是如何变化的, 并且, 如果可能, 要引进那些独立于坐标系的记号.

我们从微分流形的定义开始我们正式的讨论. 基本想法类似于我们在日常生活中所用来描述地球表面的那些东西. 人们给出一些分图表的集合, 其中每张描绘了地球的较小的相互有重叠的部分. 我们可以用观察这些分图是如何匹配的方式拼合起整张的图.

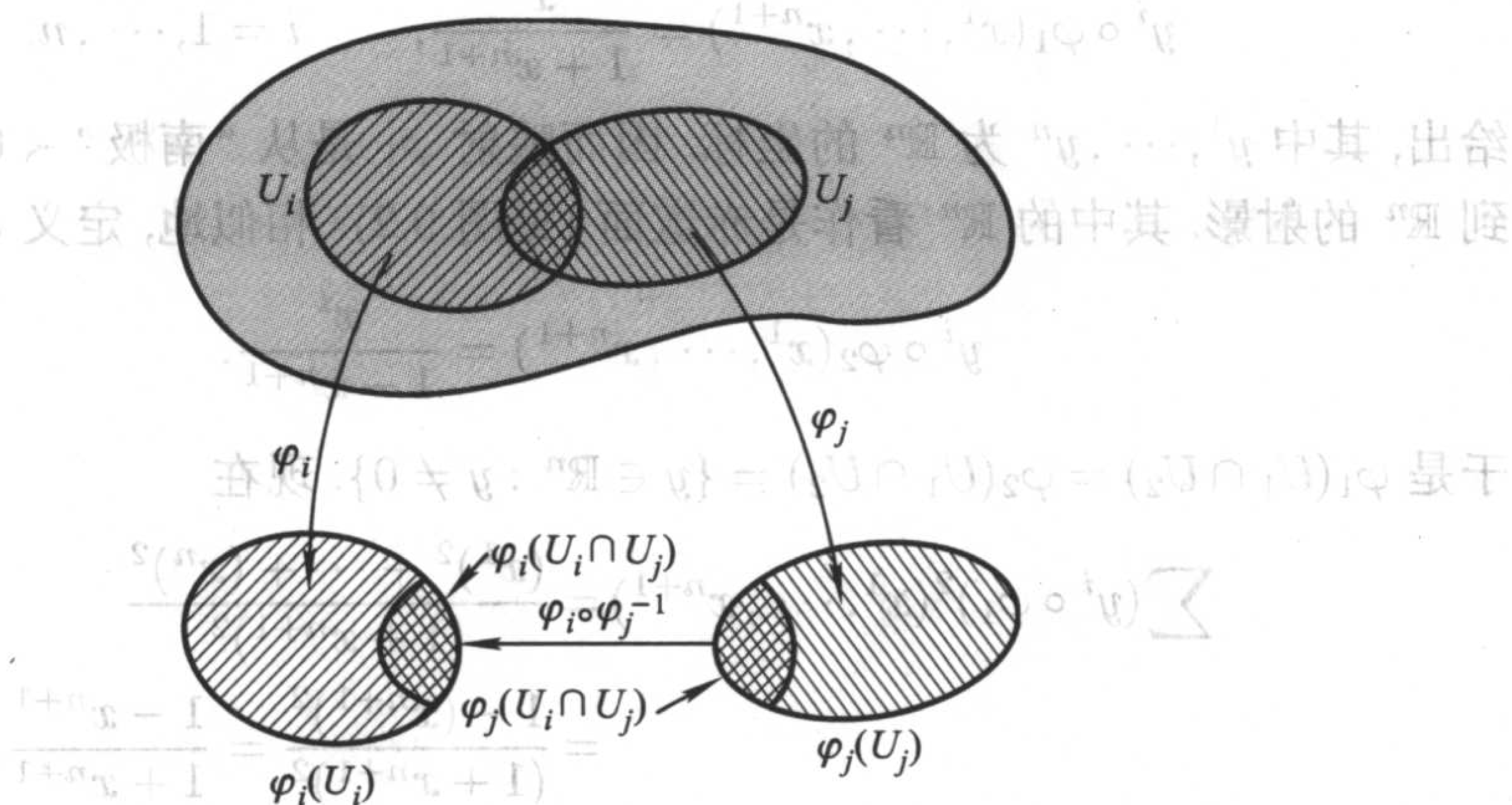


图 9.1

9.1 总图表

设 M 为一集合. V 为一巴拿赫空间.(对差不多所有的应用而言, 我们对某个整数 n 将取 V 为 $\mathbb{R}^n$.) M 上一个 C^k 类的 V - 总图表是个称作分图表的偶对 (U_i, φ_i) 的集合 $\mathfrak{a}$, 其中 U_i 为 M 的子集, φ_i 是 U_i 到 V 中一个开子集上的双射, 它们满足下列条件 (图 9.1):

A1. 对任意 $(U_i, \varphi_i) \in \mathfrak{a}$ 及 $(U_j, \varphi_j) \in \mathfrak{a}$, 集合 $\varphi_i(U_i \cap U_j)$ 和 $\varphi_j(U_i \cap U_j)$ 为 V 中开子集, 并且映射

$$\varphi_i \circ \varphi_j^{-1} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$$

是 C^k 类可微.

A2. $\cup U_i = M$.

称函数 $\varphi_i \circ \varphi_j^{-1}$ 为总图表 $\mathfrak{a}$ 的转移函数. 下面是具总图表的集合的例子.

例题 1.1 通俗例题. 设 M 为 V 中一开子集. 如果取总图表 $\mathfrak{a}$ 由单个元素 (U,φ) 组成, 其中 $U=M,\varphi:U\to V$ 为恒同映射, 则公理 A1 和 A2 毫无疑问被满足.

例题 1.2 球面. 设 $M=S^n$ 代表在 $\mathbb{R}^{n+1}$ 中由 $(x^1)^2+\cdots+(x^{n+1})^2=1$ 给出的子集. 设集合 U_1 由 $x^{n+1}>-1$ 的点构成, 而 U_2 由 $x^{n+1}<1$ 的点构成. 设

$$\varphi_1:U_1\to\mathbb{R}^n$$

由

$$y^i\circ\varphi_1(x^1,\cdots,x^{n+1})=\frac{x^i}{1+x^{n+1}},\qquad i=1,\cdots,n,$$

给出, 其中 $y^1,\cdots,y^n$ 为 $\mathbb{R}^n$ 的坐标. 因而映射 φ_1 是从"南极" $\prec 0,\cdots,0,-1\succ$ 到 $\mathbb{R}^n$ 的射影, 其中的 $\mathbb{R}^n$ 看作是赤道面 (见图 9.2). 相似地, 定义 φ_2 为

$$y^i\circ\varphi_2(x^1,\cdots,x^{n+1})=\frac{x^i}{1-x^{n+1}}.$$

于是 $\varphi_1(U_1\cap U_2)=\varphi_2(U_1\cap U_2)=\{y\in\mathbb{R}^n:y\neq 0\}$. 现在

$$\begin{aligned}\sum(y^i\circ\varphi_1)^2(x^1,\cdots,x^{n+1})&=\frac{(x^1)^2+\cdots+(x^n)^2}{(1+x^{n+1})^2}\\&=\frac{1-(x^{n+1})^2}{(1+x^{n+1})^2}=\frac{1-x^{n+1}}{1+x^{n+1}}.\end{aligned}$$

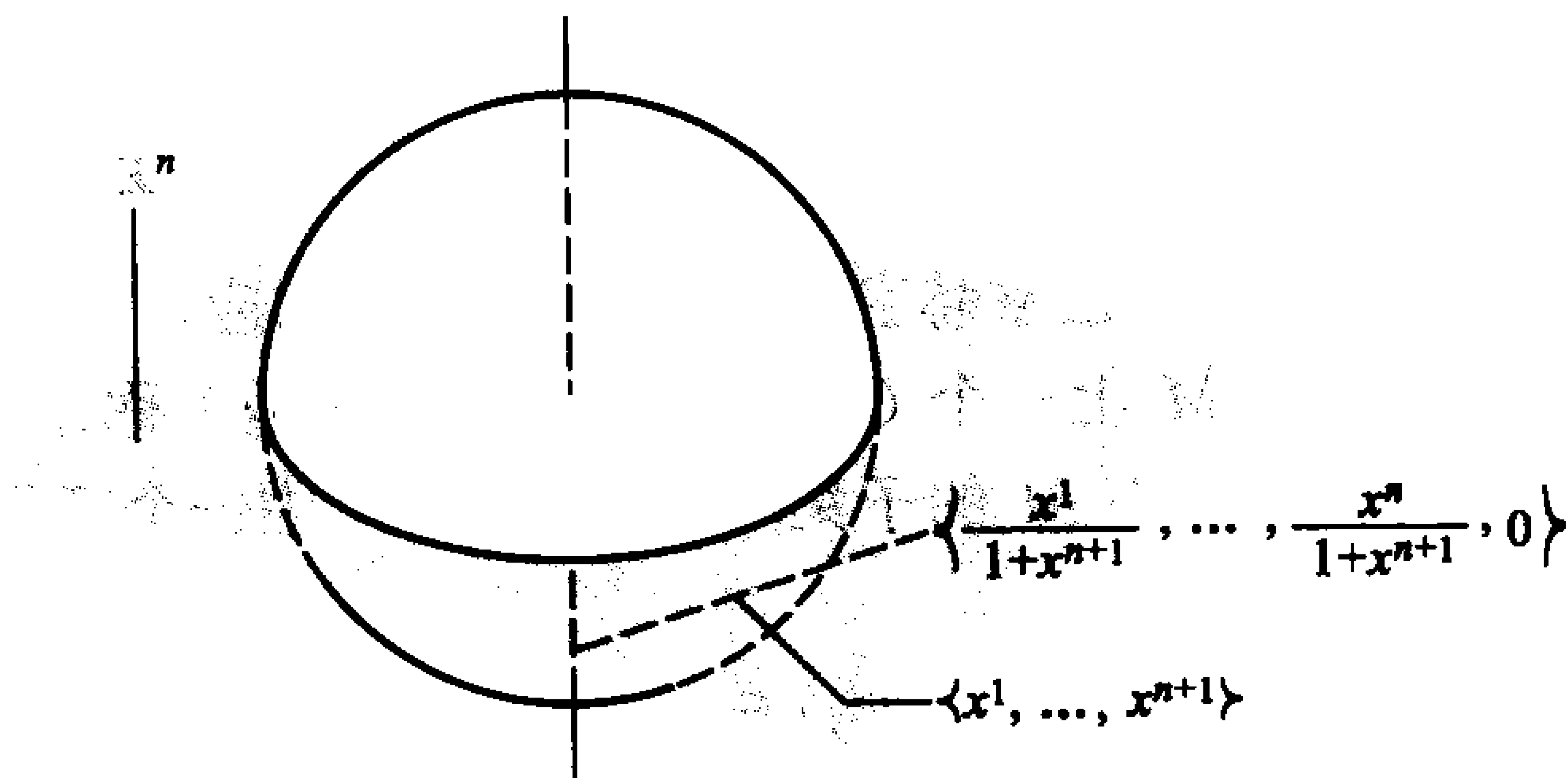


图 9.2

因此,

$$y^i \circ \varphi_2(x^1, \cdots, x^{n+1}) = \frac{y^i \circ \varphi_1(x^1, \cdots, x^{n+1})}{\sum[y^i \circ \varphi_i(x^1, \cdots, x^{n+1})]^2},$$

或者是

$$\varphi_2 = \frac{\varphi_1(x)}{||\varphi_1(x)||^2}.$$

换句话说, 映射 $\varphi_2 \circ \varphi_1^{-1}$ 定义于所有 $y \neq 0$, 并由

$$\varphi_2 \circ \varphi_1^{-1}(y) = \frac{y}{||y||^2}$$

给出. 从而条件 A1 和 A2 被满足.

注意, 我们所给出的对球面的总图表只包含了两个分图 (每个都由球极投影给出). 地球的总图表通常包含了许许多多分图. 换句话说, 许多不同的总图表可以用来描述相同的集合. 我们在后面还要回到这个论题.

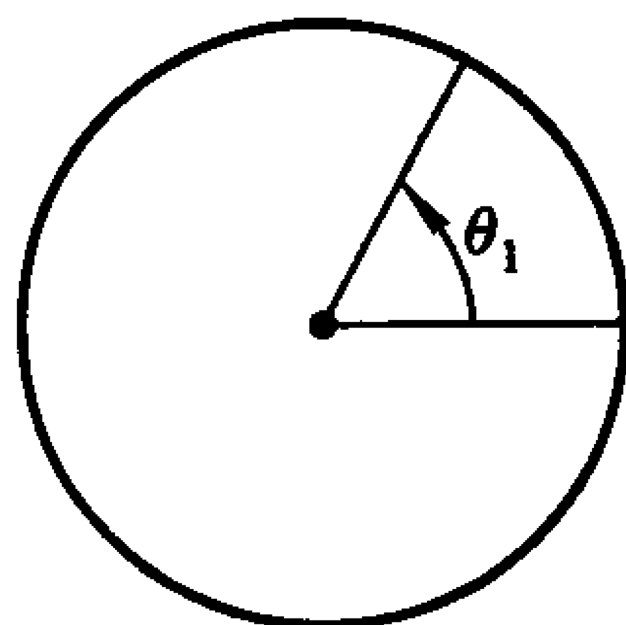


图 9.3

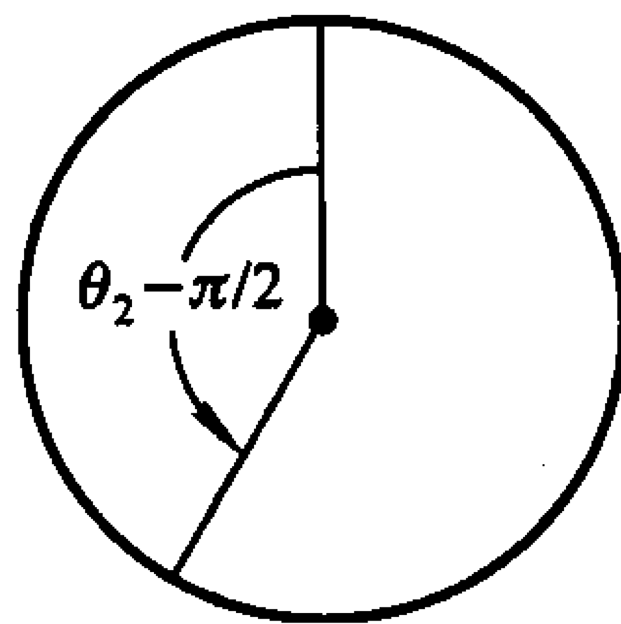


图 9.4

例题 1.3 (圆) 圆 S^1 是个 "一维球面" 从而有一个如例 1.2 中描述的总图表. 我们想以一个 S^1 上的不同总图表来描述它. 视 S^1 为单位圆 $x_1^2 + x_2^2 = 1$, 并考虑函数 θ_1, 其定义于 $\prec 1, 0 \succ$ 的一个在 S^1 上半圆的邻域, 它给出从 S^1 上这个点到 $\prec 1, 0 \succ$ 的角 (见图 9.3). 当我们反时针绕此圆运动时此函数在我们再次碰到 $\prec 1, 0 \succ$ 之前是有明确定义的. 我们要用 (U_1, θ_1) 为我们总图表中第一个分图, 其中 $U_1 = S^1 - \{\prec 1, 0 \succ\}, \theta_1$ 为上面所定义函数. 设 $U_2 = S^1 - \{\prec 0, 1 \succ\}$, 定义 θ_2 为 $\pi/2$ 加上从 $\prec 0, 1 \succ$ 到它的角 (按反时针度量)(见图 9.4). 现 $U_1 \cap U_2 = S^1 - \{\prec 1, 0 \succ, \prec 0, 1 \succ\}$, 而 $\theta_1(U_1 \cap U_2) = (0, 2\pi) - \{\pi/2\}$.



又, $\theta_2(U_1 \cap U_2) = (\pi/2, 2\pi + \pi/2) - \{2\pi\}$.

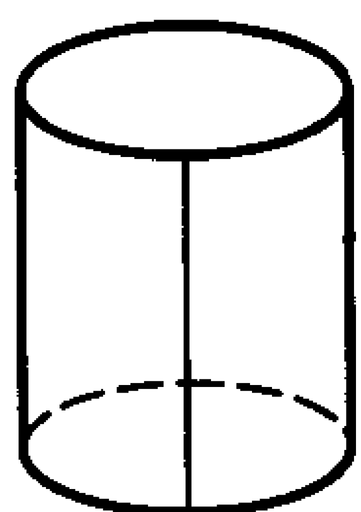


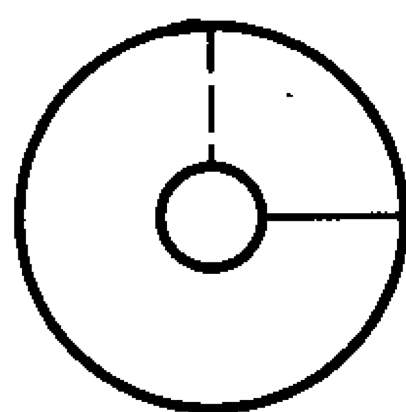
映射 $\theta_2 \circ \theta_1^{-1}$ 为

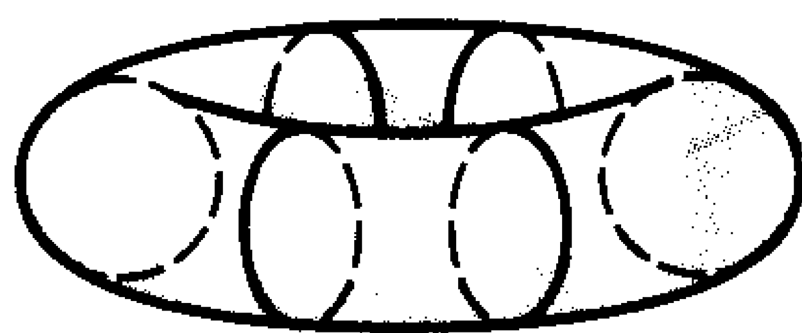
$$\theta_2 \circ \theta_1^{-1}(x) = \begin{cases} x + 2\pi, & \text{如果 } 0 < x < \pi/2, \\ x, & \text{如果 } \pi/2 < x < 2\pi. \end{cases}$$

例题 1.4　两个总图表的乘积. 设 $\mathfrak{a} = \{(U_i, \varphi_i)\}$ 为集合 M 上的一个 V_1 总图表, 又设 $\mathfrak{B} = \{(W_j, \psi_j)\}$ 为在集合 N 上的 V_2 总图表, 其中 V_1 和 V_2 都是巴拿赫空间. 则集合 $\mathfrak{C} = \{(U_i \times V_j, \varphi_i \times \psi_j\}$ 是 $M \times N$ 上的一个 $V_1 \times V_2$ 总图表. 这里的 $\varphi_i \times \psi_j(p, q) = \prec \varphi_i(p), \psi_j(q) \succ$, 其中 $\prec p, q \succ \in U_i \times W_j$. 容易验证 $\mathfrak{C}$ 满足条件 A1 和 A2. 我们称 $\mathfrak{C}$ 为 $\mathfrak{a}$ 和 $\mathfrak{B}$ 的乘积, 记为 $\mathfrak{C} = \mathfrak{a} \times \mathfrak{B}$.

举例说, 设 $M = (0, 1) \subset \mathbb{R}^1, N = S^1$. 于是我们可以将 $M \times N$ 看作一个柱面或一个圆环. 如果 $M = N = S^1$, 则 $M \times N$ 是个环面.


圆柱面


圆环


环面

写出这些情形的转移函数和总图表的显式表达式是个有意义的练习.

例题 1.5　作为第一个例子的推广, 设 S 为第三章 3.12 节定义过的一个 $(n + m)$ 维向量空间 X 的子流形. 在那里所定义的每个邻域 N, 集合 $S \cap N$ 连同映射 φ 给出了在 V 中取值的一个分图表 (这里的 X 被看成了 $V \times W$), 其中的 φ 是射影 π_1 在 S 上的限制. 在这样一个邻域 N 中, 集合 S 被表示成了函数 F 的图像. 换句话说,

$$S \cap N = \{\prec x, F(x) \succ \in V \times W : x \in \pi_1(s)\},$$

其中的 F 是从 $A = \pi_1(S \cap N)$ 到 W 中的一个光滑映射. 设 N' 是另一个这样的邻域, 其对应的射影为 π_1'(这里的 X 现在以某种其他的方式恒同于 $V \times W$). 于

是 $\varphi' \circ \varphi^{-1}(x) = \pi_1'(x, F(x))$, 这表明 $\varphi^1 \circ \varphi^{-1}$ 是个光滑映射. 因而在第三章意义下的每个子流形都具有一张总图表.

习题

设 $\mathbb{P}^n$(n 维射影空间) 表示 $\mathbb{R}^{n+1}$ 中所有通过原点的直线的空间. 任意这样一条直线被在此直线上一个非零向量所决定. 两个这样的向量 $\prec x^1, \cdots, x^n \succ$ 和 $\prec y^1, \cdots, y^{n+1} \succ$ 决定同一条直线的充要条件是它们相差一个因子, 即对所有 $i, y^i = \lambda x^i$, 其中 λ 是某个 (非零) 实数. 我们因而可将 $\mathbb{P}^n$ 中一个元素看成是非零向量的一个等价类. 对每个 i 为 1 与 $n+1$ 之间的数, 设 $U_i \subset \mathbb{P}^n$ 为由 $x^i \neq 0$ 所产生的向量的那些元素的集合. 映射

$$U_i \xrightarrow{\alpha_i} \mathbb{R}^n$$

由

$$\prec x^1, \cdots, x^{n+1} \succ \longmapsto \prec \frac{x^1}{x^i}, \cdots, \frac{x^{i-1}}{x^i}, \frac{x^{i+1}}{x^i}, \cdots, \frac{x^{n+1}}{x^i} \succ$$

给出. 证明映射 α_i 有明确定义, 而 $\{(U_i, \alpha_i)\}$ 是 $\mathbb{P}^n$ 上的一个总图表.

9.2 函数, 收敛性

设 $\mathfrak{a}$ 为集合 M 上一个 C^k 类的 V 总图表. 设 f 为定义于 M 上的一个实函数. 对每个分图表 (U_i, φ_i) 我们得到定义于 $\varphi_i(U_i)$ 上的函数 f_i, 即

$$f_i = f \circ \varphi_i^{-1}. \tag{2.1}$$

函数 f_i 可以被看作用分图表 (U_i, φ_i) 的 "f 的局部表达式". 一般来说, 这些函数 f_i 看起来彼此十分不同. 例如, 设 $M = S^n$,$\mathfrak{a}$ 为所描述过的那个总图表, 并设 f 为对每点 $\prec x^1, \cdots, x^{n+1} \succ$ 指定值 x^{n+1} 在球上的函数. 那么

$$f_1(y) = f \circ \varphi_1^{-1}(y) = \frac{2}{1 + \|y\|^2} - 1,$$

而

$$f_2(y) = f \circ \varphi_2^{-1}(y) = 1 - \frac{2}{1 + \|y\|^2},$$

这可由解出方程来验证.

回到一般性的讨论, 我们看出这些函数 f_i 并不是完全彼此无关的. 事实上, 由定义 (2.1) 式我们得到在 $\varphi_j(U_i \cap U_j)$ 上,

$$f_i \circ \varphi_i \circ \varphi_j^{-1} = f_j. \tag{2.2}$$

[因而在上面所引述的例子中我们真正有了 $f_2(y) = f_1(y/\|y\|^2)$, 这与 (2.2) 式所要求的一样.]

现在我们得到一个简单却重要的观察结果. 假设我们从一组函数 $\{f_i\}$ 出发, 其中每个 f_i 定义于 $\varphi_i(U_i)$ 上, 它们使得 (2.2) 式成立. 于是, 在 M 上便存在一个惟一的函数 f 使得 $f_i = f \circ \varphi_i^{-1}$. 事实上, 如果 $p \in U_i$, 便定义 $f(p) = f_i(\varphi_i(p))$. 要使 f 确有定义, 我们必须确信此定义是协调一致的, 即如果 p 也在 U_j 中, 则 $f_i(\varphi_i(p)) = f_j(\varphi_j(p))$, 但这正好是 (2.2) 式所讲的东西.

我们从而可按两种方式想像一个实函数: 即下面两者之一.

(i) 在 M 上以不变方式定义的一个对象, 即从 M 到 $\mathbb{R}$ 的一个映射, 或者

(ii) 一组对象 (这里是函数), 其中每一个元对每个分图表有定义并满足某些"转移法则", 即 (2.2) 式. 这种观察 M 上对象的对偶方式在后面的讨论中会相当频繁地出现.

设 M 为一具有 C^k 类总图表的集合. 我们说函数 f 是 C^l 类 ($l \leqslant k$) 的是指由 (2.1) 式所定义的函数 f_i 是 C^l 类的. 注意, 由于 $l \leqslant k$, 那么从 (2.2) 式知这不会出现任何障碍. 如果 $f_i \in C^l$ 且 $\varphi_i^{-1} \circ \varphi_j \in C^k (k \geqslant l)$, 于是 $f_j \circ (\varphi_i^{-1} \circ \varphi_j) \in C^l$. 如果 l 大于 k, 则一般地说, 如果 f_j 是 C^l 的, 那么 f_i 也不会是 C^l 类的, 这时只会有极少的 C^l 类函数.

由于我们不想一直不停地申明我们总图表的可微性次数, 从现在起, 当我们说到一张总图表时我们总是指一个类 C^∞ 的总图表.

设 M 为具总图表 $\mathfrak{a}$ 的集合. 我们关于一个点的序列 $\{x_i \in M\}$ 收敛于点 $x \in M$ 是说,

(i) 存在一个分图表 $(U_i, \varphi_i) \in \mathfrak{a}$ 和一整数 N_i, 使得 $x \in U_i$, 以及对所有 $k > N, x_k \in U_i$;

(ii) $\varphi_i(x_k)_{k>N}$ 收敛于 $\varphi(x)$.

请注意, 如果 (U_j, φ_j) 是任何一个其他的分图表使 $x \in U_j$, 则存在一个 N_j, 使得 $\varphi_j(x_k) \in U_j$ 对所有 $k > N_j$ 成立, 并且 $\varphi_j(x_k) \to \varphi_j(x)$. 事实上, 选取 N_j 使得 $\varphi_i(x_k) \in \varphi_i(U_i \cap U_j)$ 对所有 $k \geqslant N_j$ 成立即可.(由于 A1, $\varphi_j(U_i \cap U_j)$ 是开集, 故这是可能的.) $\varphi_j(x_k)$ 收敛于 $\varphi_j(x)$ 的事实来自 $\varphi_j \circ \varphi_i^{-1}$ 的连续性.因此说 $\{x_i\}$ 收敛于 x 是有完全明确的意义.

警示　说序列 $\{x_i\}$ 是柯西的则没有意义. 举例来说, 设 $M = S^n$, 其总图表如前面所描述的. 如果 $\{x_k\}$ 是 S^n 上收敛于北极点的点序列, 则 $\varphi_1(x_k) \to 0$ 而 $\varphi_2(x_k) \to \infty$. 如果我们去掉北极点, 即 $M = S^n - \{\prec 0, \cdots, 0, 1 \succ\}$, 并定义这些

图表仍如前, 则这个例子更加棘手. 于是 $\{x_k\}$ 在 M 中没有极限. 显然 $\{\varphi_1(x_k)\}$ 是柯西序列而 $\{\varphi_2(x_k)\}$ 则不是的.

一旦我们有了收敛性的概念, 我们则可谈及诸如开集和闭集之类的东西. 我们也可以直接定义. 譬如说, 一个集合 U 是开的如果 $\varphi_i(U\cap U_i)$ 是 $\varphi_i(U_i)$ 的开子集合, 其中 (U_i,φ_i) 为所有的分图表, 等等.

习题

2.1 证明上面对集合为开的定义是前后一致的, 即存在有非空开集.(事实上, 每个 U_i 都是开的.)

2.2 证明序列 $\{x_\alpha\}$ 收敛于 x 当且仅当对每个包含 x 的开集 U, 存在一个 N_U, 使得对 $\alpha > N_U$ 有 $x_\alpha \in U$.

设 $\mathfrak{a}=\{(U_i,\varphi_i)\}$ 为 M 上一个总图表, U 为 M 的一个开子集. 设 $\mathfrak{a}\restriction U$ 为所有偶对 $(U_i\cap U,\varphi_i\restriction U)$ 的集合. 容易验证 $\mathfrak{a}\restriction U$ 是 U 上的一个总图表. 称其为 $\mathfrak{a}$ 在 U 上的限制.

设 f 是定义于开集 U 上的函数. 我们称 f 在 U 上是 C^l 类的是说它关于在 U 上的图表 $\mathfrak{a}\restriction U$ 是 C^l 类的. 为了后面使用便利, 我们称定义于 M 的一个子集上的函数 f 是 C^l 类的是说, 如果

(i) f 的定义域是 M 的某个开集, 并且

(ii) f 在 U 上为 C^l 类.

9.3 微分流形

在我们第 1 节所讨论的例子中, 我们在每种情形中所做的对总图表的个别选取是十分随意的. 我们能完全等价地引进不同的总图表而不改变可微函数类, 或者开集的类, 或者收敛序列, 等等. 因此, 我们在 M 的总图表间引进一个等价关系:

设 $\mathfrak{a}_1$ 和 $\mathfrak{a}_2$ 为 M 上的两个总图表. 我们说这两个是等价的指的是如果它们的并 $\mathfrak{a}_1\cup\mathfrak{a}_2$ 仍是 M 上的一个总图表.

关键的条件是 A1 对这个并集仍然成立. 这意味着对任何分图表 $(U_i,\varphi_i)\in\mathfrak{a}_1$ 和 $(W_j,\psi_j)\in\mathfrak{a}_2$, 集合 $\varphi_i(U_i\cap W_j)$ 和 $\psi_j(U_i\cap W_j)$ 是开集, 并且 $\varphi_i\circ\psi_j^{-1}$ 是 $\psi_j(U_i\cap W_j)$ 到 $\varphi_i(U_i\cap W_j)$ 上的可微映射, 同时具有可微逆.

所引进的关系很清楚是个等价关系. 进一步, 作为一个易做的习题可验证, 如果 f 关于一个给定的总图表是 C^l 类的, 则它关于任一个等价的图表也是 C^l 类的. 对于开集和收敛性的概念也同样如此.

定义 3.1 称集合 M 连同 M 上的一个总图表的等价类为一个微分流形是说它满足"豪斯多夫性质": 对 M 的任意两个点 $x_1 \neq x_2$, 存在开集 U_1 和 U_2 使 $x_1 \in U_1, x_2 \in U_2$ 且 $U_1 \cap U_2 = \phi$.

在下文中我们仍以 M 代表 (不恰当用语) 一个微分流形, 其中的总图表等价类不言自明. M 的一个总图表的意思是属于这个给定的等价类中的一个图表, 而 M 的一个分图表是表示属于 M 某总图表中的一个分图表.

我们仍然采用这样的记号约定, 即 V 是 M 上分图表取值的巴拿赫空间 (我们这时说 M 是个 V - 流形). 如果有多个流形, M_1, M_2 等等被论及, 我们将以 V_1, V_2 等表示相应的向量空间. 如果 $V = \mathbb{R}^n$, 我们则说 M 是个 n维流形.

设 M_1 和 M_2 为微分流形. 映射 $\varphi : M_1 \to M_2$ 是连续的指的是对任意开集 $U_2 \subset M_2$, 集合 $\varphi^{-1}(U_2)$ 为 M_1 的开子集. 设 $x_2 \in M_2, U_2$ 为任一包含 x_2 的开集. 如果 $\varphi(x_1) = x_2$, 则 $\varphi^{-1}(U_2)$ 是包含 x_1 的一个开集. 如果 (W, α) 是一个关于 x_1 的分图表, 则 $W \cap \varphi^{-1}(U_2)$ 是 W 的一个开子集, 而 $\alpha(W \cap \varphi^{-1}(U_2))$ 是 V_1 中包含 $\alpha(x_1)$ 的开集. 因此存在一个 $\varepsilon > 0$, 使得对所有 $x \in W$, 满足 $\|\alpha(x) - \alpha(x_1)\| < \varepsilon$ 有 $\varphi(x) \in U_2$. 在这种意义下可以说, 所有"靠近 x_1"的点被映到"靠近 x_2". 注意, ε 的选取要同时依赖于分图表 (W, α) 和 x_1, x_2, U_2 以及 φ.

如果 M_1, M_2 和 M_3 为微分流形, 并且如果 $\varphi : M_1 \to M_2$ 和 $\psi : M_2 \to M_3$ 为连续映射, 则容易看出它们的复合 $\psi \circ \varphi$ 是从 M_1 到 M_3 的连续映射.

设 φ 为 M_1 到 M_2 的连续映射. 设 (W_1, α_1) 为 M_1 上的分图表, (W_2, α_2) 为 M_2 上的分图表. 我们称这些分图表是 (在 φ 下)相容的是指如果有 $\varphi(W_1) \subset W_2$. 如果 $\mathfrak{a}_2$ 是 M_2 上的总图表, $\mathfrak{a}_1$ 是 M_1 上的总图表, 我们说 $\mathfrak{a}_1$ 与 $\mathfrak{a}_2$ 在 φ 下是相容的是指如果对每个 $(W_1, \alpha_1) \in \mathfrak{a}_1$ 存在一个 $(W_2, \alpha_2) \in \mathfrak{a}_2$ 与它相容, 即使得 $\varphi(W_1) \subset W_2$.(注意, 映射 $\alpha_2 \circ (\varphi \restriction W_1) \circ \alpha_1^{-1}$, 于是是一个从 V_1 中一个开子集到 V_2 中的一个连续映射.) 给出了 $\mathfrak{a}_2$ 和 φ, 我们总能找到一个 $\mathfrak{a}_1$ 使它在 φ 下与 $\mathfrak{a}_2$ 相容. 事实上, 设 $\mathfrak{a}_1'$ 为 M_1 上任一个总图表, 并令

$$\mathfrak{a}_1 = \{(W_1 \cap \varphi^{-1}(W_2), \alpha \restriction (W_1 \cap \varphi^{-1}(W_2))\},$$

其中 (W_1, α) 遍历 $\mathfrak{a}_1'$ 的所有分图表, 而 (W_2, β) 遍历所有 $\mathfrak{a}_2$ 的分图表.

定义 3.2 设 M_1 和 M_2 为微分流形, φ 为映射: $M_1 \xrightarrow{\varphi} M_2$. 称 φ 为可微的是说如果下面的条件成立:

(i) φ 连续.

(ii) 如果 $\mathfrak{a}_1$ 和 $\mathfrak{a}_2$ 为在 φ 下相容的总图表. 于是对任意相容的 $(W_1,\alpha_1) \in \mathfrak{a}_1$ 和 $(W_2,\alpha_2) \in \mathfrak{a}_2$, 映射

$$\alpha_2 \circ \varphi \circ \alpha_1^{-1} : \alpha_1(W_1) \to \alpha_2(W_2)$$

是可微的 (作为从一个巴拿赫空间中开子集到一个巴拿赫空间的映射)(见图 9.5).

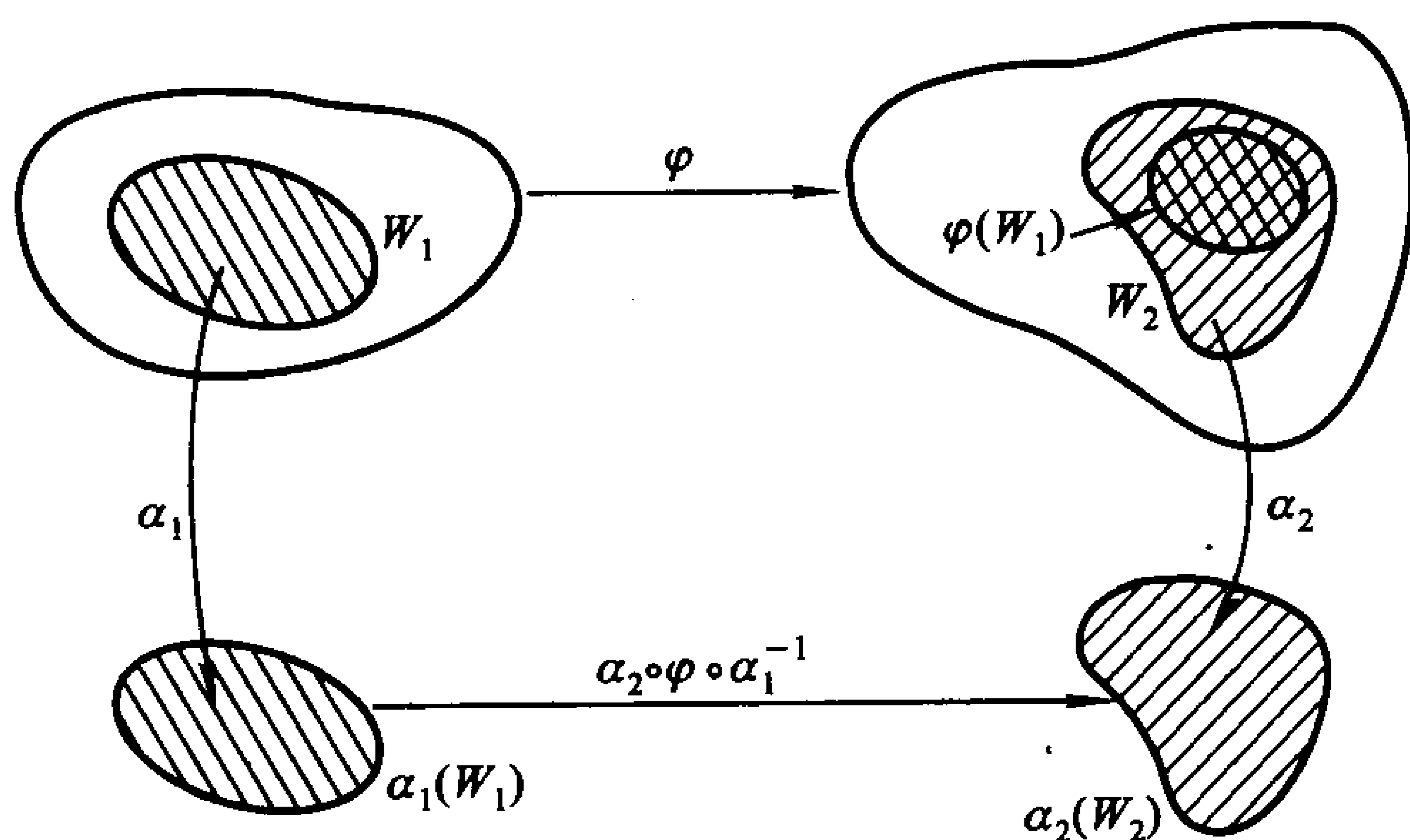


图 9.5

为了验证一个连续映射 φ 是可微的, 验证远比 (ii) 所要求的少得多的东西就够了. 条件 (ii) 所关联到的是任何一对相容的总图表和任何一对相容的分图表. 事实上, 我们可以断言:

命题 3.1 设 $\varphi: M_1 \to M_2$ 连续, $\mathfrak{a}_1$ 和 $\mathfrak{a}_2$ 为在 φ 下相容的总图表. 假设对每个 $(W_1,\alpha_1) \in \mathfrak{a}_1$ 存在一个 $(W_2,\alpha_2) \in \mathfrak{a}_2$, 满足 $\varphi(W_1) \subset W_2$ 和 $\alpha_2 \circ \varphi \circ \alpha_1^{-1}$ 可微. 则 φ 可微.

证明 设 (U_1,β_1) 和 (U_2,β_2) 为 M_1, M_2 上任意分图表, 满足 $\varphi(U_1) \subset U_2$. 我们必须证明 $\beta_2 \circ \varphi \circ \beta_1^{-1}$ 可微. 只要证明它在每个点 $\beta_1(x_1), x_1 \in U_1$ 的邻域中可微. 选取 $(W_1,\alpha_1)\in \mathfrak{a}_1, x \in W_1$, 又选取 $(W_2,\alpha_1) \in \mathfrak{a}_2$, 满足 $\varphi(W_1) \subset W_2$. 于是在 $\beta_1(W_1 \cap U_1)$ 上我们有

$$\beta_2 \circ \varphi \circ \beta_1^{-1} = (\beta_2 \circ \alpha_2^{-1}) \circ (\alpha_2 \circ \varphi \circ \alpha_1^{-1}) \circ (\alpha_1 \circ \beta_1^{-1}),$$

故左端是可微的. □

换句话说, 只要对一对总图表证实其可微性就可以了. 作为推论我们有

命题 3.2 设 $\varphi: M_1 \to M_2, \psi: M_2 \to M_3$ 可微, 则 $\psi \circ \varphi$ 可微.

证明 设 $\mathfrak{a}_3$ 为 M_3 的一个总图表. 选取 $\mathfrak{a}_2$ 与 $\mathfrak{a}_3$ 在 ψ 下相容, 然后再选取 M_1 上一个总图表 $\mathfrak{a}_1$ 在 φ 下与 $\mathfrak{a}_2$ 相容. 对任意 $(W_1, \alpha_1) \in \mathfrak{a}_1$, 选取 $(W_2, \alpha_2) \in \mathfrak{a}_2, (W_3, \alpha_3) \in \mathfrak{a}_3$, 满足 $\varphi(W_1) \subset W_2$ 和 $\psi(W_2) \subset W_3$. 于是 $\alpha_3 \circ \psi \circ \varphi \circ \alpha_1^{-1} = (\alpha_3 \circ \psi \circ \alpha_2^{-1}) \circ (\alpha_2 \circ \varphi \circ \alpha_1^{-1})$ 可微. □

习题

3.1 设 $M_1 = S^n, M_2 = \mathbb{P}^n$, 以及 $\varphi: M_1 \to M_2$ 为将单位球上每点映到它确定的直线.(注意, S^n 的两个对经点变到 $\mathbb{P}^n$ 的同一点.) 构造 φ 的一组相容总图表并证明 φ 可微.

我们留意到, 如果 f 是 M 上任意的取值于巴拿赫空间的函数, 则 f 作为一个函数是可微的 (在第 2 节的意义下) 当且仅当它作为流形间的映射是可微的. 特别, 设 $\varphi: M_1 \to M_2$ 为可微映射,f 为 M_2 上的可微函数 (定义于某个开子集上, 譬如 U_2). 于是 $f \circ \varphi$ 便是 M_1 上的可微函数 [定义于开集 $\varphi^{-1}(U_2)$ 上]. 因此 φ 把 M_2 上一个可微函数“拉回”到 M_1. 由此观点, 我们可以说 φ 诱导了从 M_2 上的可微函数集到 M_1 上可微函数集的一个映射. 我们将此映射记为 φ^*. 因而

$$M_2\text{上的可微函数} \xrightarrow{\varphi^*} M_1\text{上的可微函数},$$

它由

$$\varphi^*[f] = f \circ \varphi$$

给出. 如果 $\psi: M_2 \to M_3$ 为第二个可微映射, 则 $(\psi \circ \varphi)^*$ 从 M_3 上的函数走到 M_1 上的可微函数, 并且有

$$(\psi \circ \varphi)^* = \varphi^* \circ \psi^* \tag{3.1}$$

(注意顺序的变化). 事实上, 对 M_3 上的 g, 有

$$(\psi \circ \varphi)^* g = g \circ (\psi \circ \varphi) = (g \circ \psi) \circ \varphi = \varphi^*[\psi^*[g]].$$

我们看出, 如果 φ 是 $M_1 \to M_2$ 的任意映射, 而 f 为定义在 M_2 上任意子集 S_2 上的任意函数, 则“拉回” $\varphi^*[f] = f \circ \varphi$ 是定义在 M_1 的 $\varphi^{-1}(S_2)$ 上的函数.

φ 为连续这个事实让我们得到结论说, 如果 S_2 为开, 则 $\varphi^{-1}(S_2)$ 也是. φ 是可微的这个事实蕴涵了一旦 f 是可微的则 $\varphi^*[f]$ 也是的.

一旦代数运算有定义, 则映射 φ^* 与所有这些代数运算可交换. 更准确地说, 假设 f 和 g 取值于同一向量空间, 并有定义域 U_1 和 U_2, 则 $f+g$ 在 $U_1 \cap U_2$ 上有定义, 并且 $\varphi^*[f]+\varphi^*[g]$ 在 $\varphi^{-1}(U_1 \cap U_2)$ 上有定义, 显然有

$$\varphi^*[f+g] = \varphi^*[f] + \varphi^*[g].$$

习题

3.2 设 M_2 为有限维流形, $\varphi : M_1 \to M_2$ 连续. 假设 $\varphi^*[f]$ 对任意 (局部定义的) 可微实函数 f 可微. 得出 φ 是可微的结论.

3.3 证明如果 φ 是巴拿赫空间的有界线性映射, 则上面定义的 φ^* 是第二章第 3 节定义的 φ^* 的一个扩张.

9.4 切空间

在本节中我们要对一个微分流形在此流形的每点上构造一个“近似的向量空间”. 这将使我们能详细叙述在流形上的微分学的大多数概念.

设 M 是个微分流形, x 为 M 的一点 (图 9.6). 设 $I \subset \mathbb{R}$ 为包含原点的区间设 φ 为 I 到 M 的一个可微映射使得 $\varphi(0) = x$. 我们称 φ 为一条通过 x 的 (可微) 曲线.

设 f 为 M 上任意可微实函数, 定义于 x 的一个邻域. 于是 $\varphi^*[f]$ 在 $\mathbb{R}$ 上可微, 并且我们可以考虑它在原点的导数. 定义算子 D_φ 为

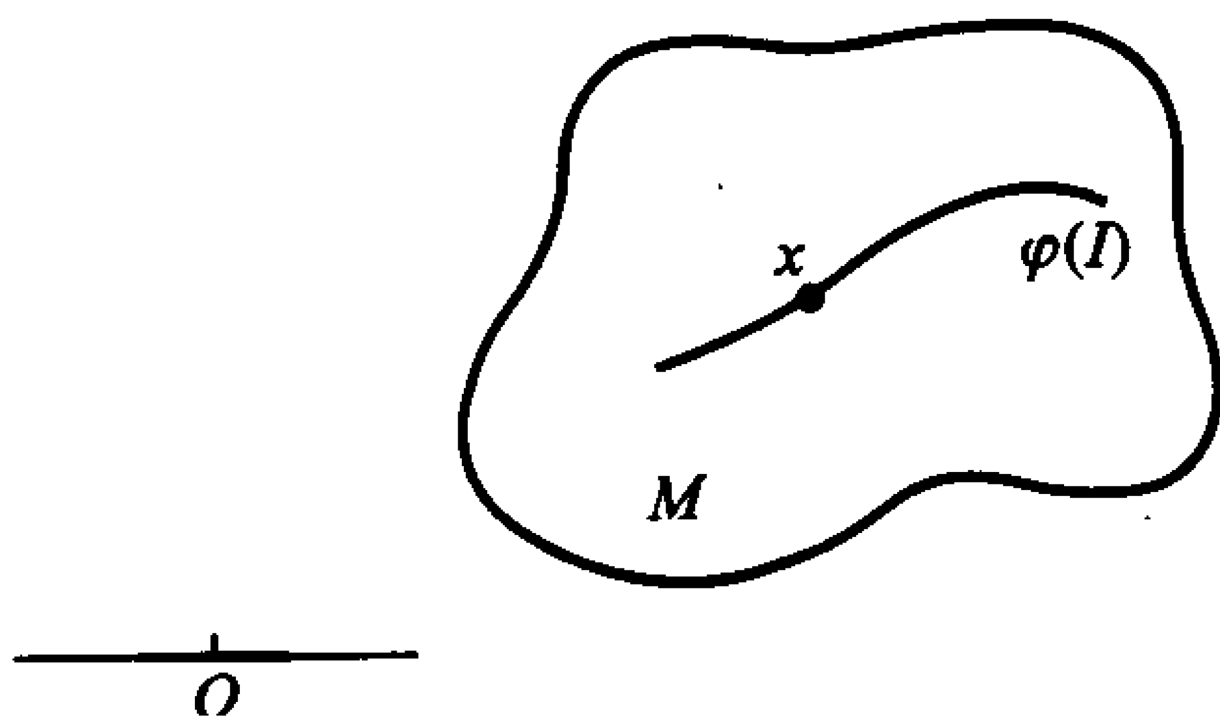


图 9.6

$$D_\varphi(f) = \left.\frac{d\varphi^*[f]}{dt}\right|_{t=0}.$$

由于 φ^* 的线性性, 映射 $f \mapsto D_\varphi(f)$ 是线性的:

$$D_\varphi(af+bg) = aD_\varphi(f) + bD_\varphi(g).$$

相似地，我们有莱布尼茨规则

$$D_\varphi(fg) = f(x)D_\varphi(g) + g(x)D_\varphi(f),$$

容易验证它. 泛函 D_φ 依赖于曲线 φ. 如果 ψ 是第二条曲线，则一般地，$D_\varphi \neq D_\psi$. 然而如果 $D_\psi = D_\varphi$，则我们说曲线 φ 与 ψ 在 x 相切，并记为 $\varphi \sim \psi$. 因此 $\varphi \sim \psi$ 当且仅当 $D_\varphi(f) = D_\psi(f)$ 对所有可微函数 f 成立.

容易验证 $\sim$ 是个等价关系. 通过 x 的曲线的一个等价类被称做在 x 的一个切向量. 如果 ξ 是在 x 的一个切向量，$\varphi \in \xi$，我们则说 ξ 在 x 切于 φ.

对任何在 x 附近有定义的可微函数 f 和任意切向量 ξ，我们令

$$\xi(f) = D_\varphi(f),$$

其中 $\varphi \in \xi$. 因而 ξ 给出了在 x 附近定义的可微函数上的一个泛函. 我们有

$$\xi(af + bg) = a\xi(f) + b\xi(g), \tag{4.1}$$

$$\xi(fg) = f(x)\xi(g) + g(x)\xi(f). \tag{4.2}$$

让我们来考查一下，用 x 附近的分图表 (W, α) 来表示，等价关系 $\sim$ 到底说了些什么. 泛函 $D_\varphi(f)$ 可以写为

$$\left.\frac{df \circ \varphi}{dt}\right|_{t=0} = \left.\frac{d(f \circ \alpha^{-1}) \circ (\alpha \circ \varphi)}{dt}\right|_{t=0},$$

如果令 $\Phi = \alpha \circ \varphi, F = f \circ \alpha^{-1}$，则 Φ 是一个巴拿赫空间中的参数曲线，而 F 是那里的一个可微函数. 因而我们可以写成

$$D_\varphi(f) = dF(\Phi'(0)) = D_{\Phi'(0)}F.$$

由这个表达式我们看出 (令 $\Psi = \alpha \circ \psi$)$\psi \backsim \varphi$ 当且仅当 $\Phi'(0) = \Psi'(0)$. 因而我们看出以分图表 (W, α) 表示时，在 x 的每个切向量 ξ 对应于一个惟一的向量 $\xi_\alpha \in V$，由

$$\xi_\alpha = (\alpha \circ \varphi)'(0)$$

给出，其中 $\varphi \in \xi$.

反之，给出任一 $v \in V$，有一个切向量 ξ，满足 $\xi_\alpha = v$. 事实上，定义 φ 为 $\varphi(t) = \alpha^{-1}(\alpha(x) + tv)$. 于是 φ 在 O 的一个足够小的区间有定义，并且 $(\alpha \circ \varphi)' = v$.

简而言之, 选取分图表让我们把在 x 的所有切向量与 V 等同. 设 (U,β) 为 x 附近的第二个分图表. 则

$$\xi_\beta = (\beta \circ \varphi)'(0) = [(\beta \circ \alpha^{-1}) \circ (\alpha \circ \varphi)]'(0).$$

由链规则我们得到

$$\xi_\beta = J_{\beta\circ\alpha^{-1}}(\alpha(x))\xi_\alpha, \tag{4.3}$$

其中 $J_\gamma(p)$ 是 γ 在 p 的微分 $d\gamma_p$.

由于 $J_{\beta\circ\alpha^{-1}}(\alpha(x))$ 是 V 到自己的线性映射, 所以等式 (4.3) 说在 x 的所有切向量的集合可以等同于 V, 而此等同是在 V 的一个自同构上被决定的. 特别, 我们可以把在 x 的所有切向量的集合做成一个向量空间, 其中定义

$$a\xi + b\eta = \zeta,$$

这里的 ζ 由

$$a\xi_\alpha + b\eta_\alpha = \zeta_\alpha$$

决定, α 为某个分图表. 等式 (4.3) 表明此定义与 α 无关.

我们表示这个在 x 的切向量的空间为 $T_x(M)$, 并称它为 (M) 在 x 的切空间.

设 ψ 为 M_1 到 M_2 的可微映射, 并设 φ 为通过 $x \in M_1$ 的一条曲线 (见图 9.7), 于是 $\psi \circ \varphi$ 为通过 $\psi(x) \in M_2$ 的一条曲线. 容易验证, 如果 $\varphi \backsim \overline{\varphi}$ 则 $\psi \circ \varphi \backsim \psi \circ \overline{\varphi}$. 因此映射 ψ 诱导了一个从 $T_x(M_1)$ 到 $T_{\psi(x)}(M_2)$ 的映射. 我们将以 ψ_{*x} 表示它. 重复说一遍, 如果 $\xi \in T_x(M_1)$, 则 $\psi_{*x}(\xi) = \eta$ 由对所有 $\varphi \in \xi$,

$$\psi \circ \varphi \in \eta$$

所决定.

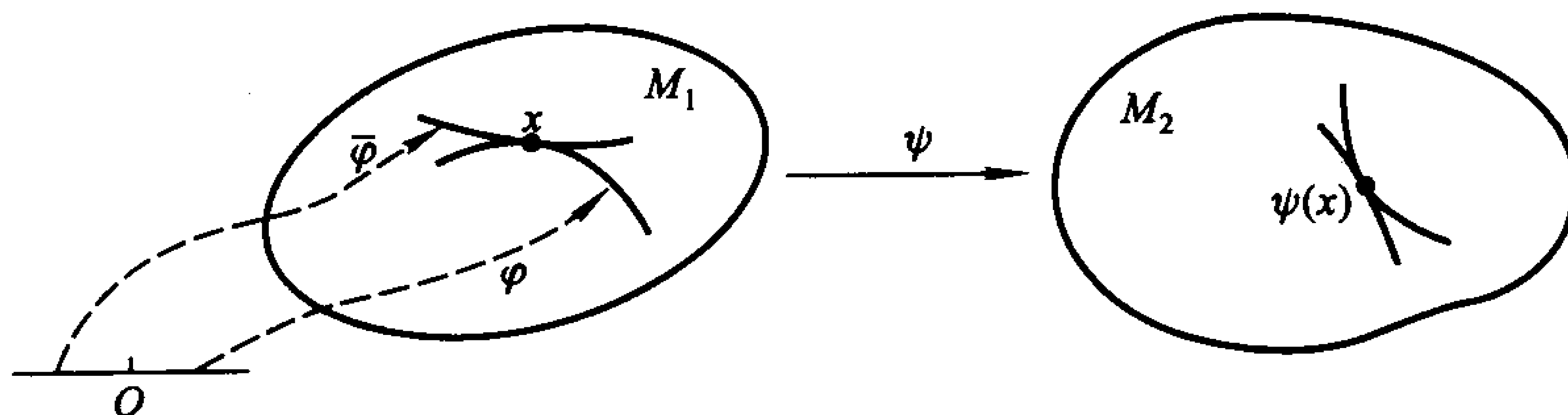


图 9.7

设 (U,α) 是 x 的一个分图表, (W,β) 是 $\psi(x)$ 的一个分图表. 于是

$$\xi_\alpha = (\alpha \circ \varphi)'(0)$$

以及

$$\eta_\beta = (\beta \circ \psi \circ \varphi)'(0) = [(\beta \circ \psi \circ \alpha^{-1}) \circ (\alpha \circ \varphi)]'(0).$$

由链规则我们因而可写成

$$\eta_\beta = J_{\beta\circ\psi\circ\alpha^{-1}}(\alpha(x))\xi_\alpha.$$

这说明如果我们通过 α 将 $T_x(M_1)$ 与 V_1 等同, $T_{\psi(x)}(M_2)$ 通过 β 与 V_2 等同, 则 ψ_{*x} 成为与线性映射 $J_{\beta\circ\psi\circ\alpha^{-1}}(\alpha(x))$ 等同的映射. 特别, 映射 ψ_{*x} 是从 $T_x(M_1)$ 到 $T_{\psi(x)}(M_2)$ 的**连续线性映射**. 如果 $\varphi: M_1 \to M_2, \psi: M_2 \to M_3$ 为两个可微映射, 则立刻由定义得到

$$(\psi \circ \varphi)_{*x} = \psi_{*\varphi(x)} \circ \varphi_{*x}. \tag{4.4}$$

我们已经看到, 分图表 (U,α) 的选取将 $T_x(M)$ 等同于 V. 现假设 M 实际上就是 V 自己 (或 V 的一个开子集), 它被看成是一个微分流形. 于是 M 有一个特别的分图表, 即 (M, id). 因而在 V 的开子集上, 这个等同分图表给了我们一种特别的方式将 $T_x(M)$ 等同于 V. 有时将 $T_x(M)$ 描绘为将原点移至 x 的一个 V 的翻版比较方便. 我们于是可将在 x 的一个切向量画成从 x 出发的一个箭头 (见图 9.8).

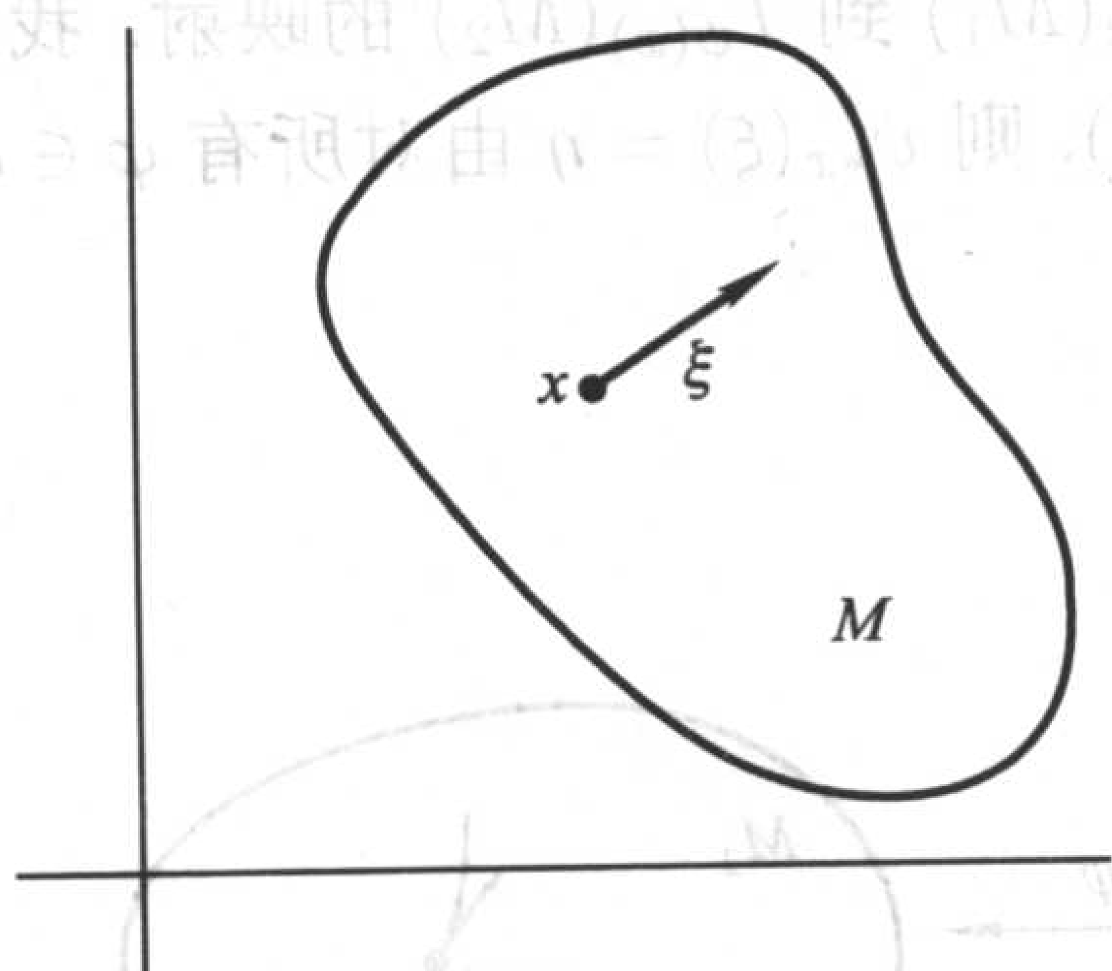


图 9.8

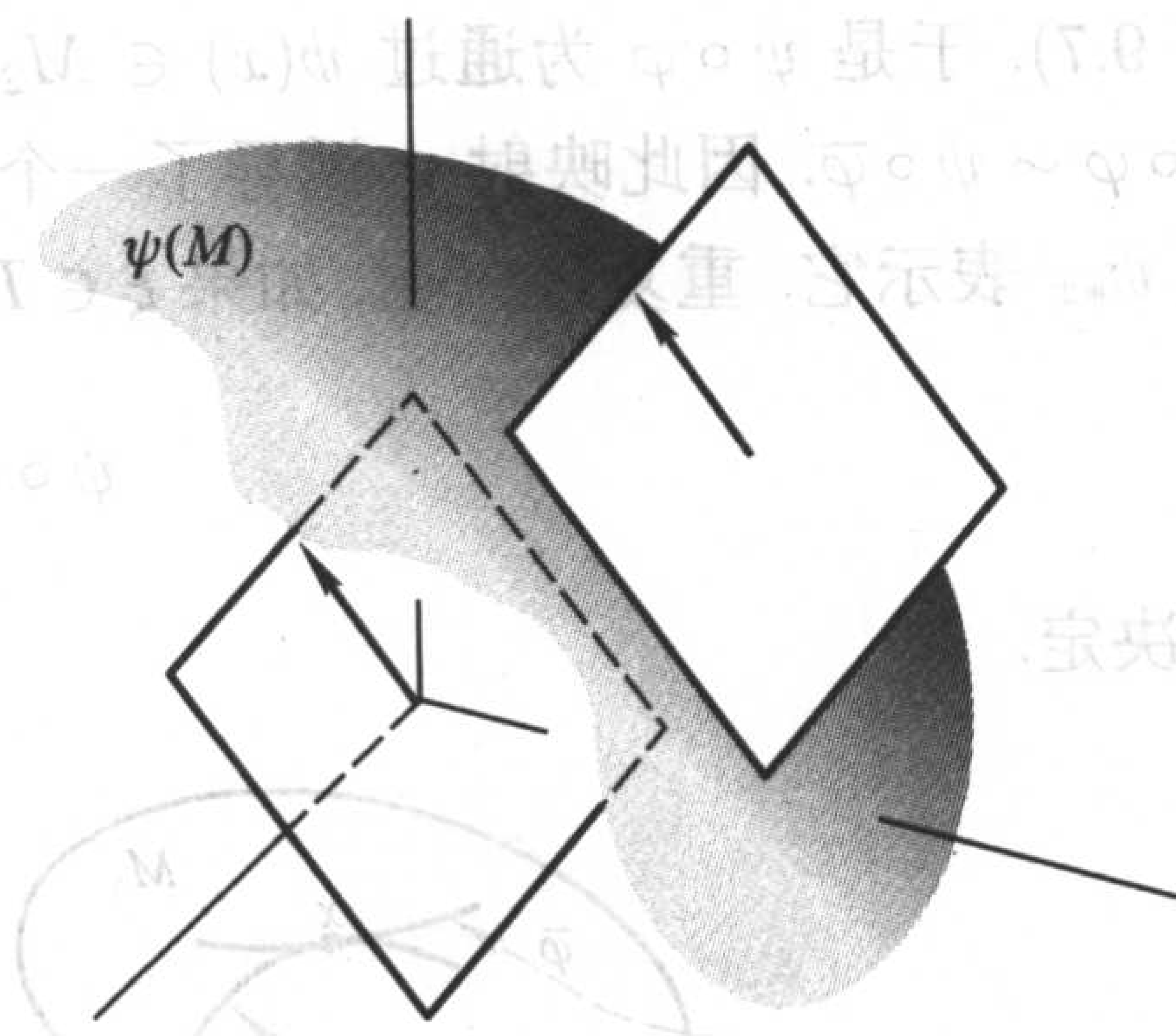


图 9.9

现假设 M 是个一般的流形, ψ 是 M 到一个向量空间 V_1 的一个微分映射. 于是 $\psi_*(T_x(M))$ 是 $T_{\psi(x)}(V_1)$ 的子空间. 如果我们将 $\psi_*(T_x(M))$ 视为 V_1 的一个

子空间, 并考虑相应的通过 x 的超平面, 我们便得到了 "在 x 切于 $\psi(M)$ 的平面", 这就是直观意义下的切平面 (图 9.9).

按这种方式来想像切向量非常方便, 那就是说, 如果 M 被映到一个向量空间, 就把它们看做 M 的切向量.

如果 f 为定义在 $x \in M$ 的一个邻域 U 中的一个实可微函数, 我们则可以将其视为从流形 U 到流形 $\mathbb{R}^1$ 的一个映射. 因此我们得到一个映射 $f_{*x}: T_x(M) \to T_{f(x)}(\mathbb{R}^1)$. 回想起我们把 $T_y(\mathbb{R}^1)$ 对任意 $y \in \mathbb{R}^1$ 等同于 $\mathbb{R}^1$. 那么, f_{*x} 便可以被看作从 $T_{*x}(M)$ 到 $\mathbb{R}^1$ 的一个映射. 读者应该验证一下, 这个映射的确由

$$f_{*x}(\xi) = \xi(f) \tag{4.5}$$

给出, 其中 $\xi \in T_x(M)$.

特别, 如果我们在 (4.4) 式中取 $M_3 = \mathbb{R}, \psi = f$, 我们便能断定:

设 ψ 是 M_1 到 M_2 的可微映射, f 为 M_2 上的可微函数, 它定义于 $\psi(x)$ 的一个邻域中. 则对任意 $\xi \in T_x(M_1)$ 有

$$\xi(\psi^*(f)) = \psi_{*x}(\xi)(f). \tag{4.6}$$

从现在起, 我们上下文可自明的情形下将常常去掉 ψ_{*x} 中的下标 x. 因而 (4.4) 式会被写为 $(\psi \circ \varphi)_* = \psi_* \circ \varphi_*$. 某些作者称映射 ψ_{*x} 为 ψ 在 x 的微分并以 $d\psi_x$ 表示它. 如果 M_1 和 M_2 分别为巴拿赫空间 V_1 和 V_2 的开子集 (因而在它们的等同分图表下为微分流形.), 则上面所定义的 ψ_{*x} 当 $T_x(M_i)$ 等同于 V_i 时便约化为微分 $d\psi_x$. 这个约化依赖于这个等同表示.

9.5 流与向量场

设 M_1 和 M_2 为微分流形. 映射 $g: M_1 \to M_2$ 被称作一个微分同胚是说, 如果 g 是从 M_1 到 M_2 上的可微的一一映射, 使得 g^{-1} 也是可微的.

设 M 为微分流形. 映射 $\varphi: M \times \mathbb{R} \to M$ 是个单参数群说的是, 如果

(i) φ 可微;

(ii) $\varphi(x, 0) = x$ 对所有 $x \in M$ 成立;

(iii) $\varphi(\varphi(x, s), t) = \varphi(x, s+t)$ 对所有 $x \in M$ 和 $s, t \in \mathbb{R}$ 成立.

我们可以以稍稍不同的方式来表达条件 (ii) 和 (iii). 设 $\varphi_t: M \to M$ 由

$$\varphi_t(x) = \varphi(x, t).$$

对每个 $t \in \mathbb{R}$, 映射 φ_t 为可微. 事实上,

$$\varphi_t = \varphi \circ \iota_t,$$

其中 ι_t 是由 $\iota_t(x) = (x, t)$ 给出的 $M \to M \times \mathbb{R}$ 的可微映射.

条件 (ii) 说的是 φ_0=id. 条件 (iii) 说的是

$$\varphi_t \circ \varphi_s = \varphi_{t+s}.$$

如果我们取 $t = -s$ 于上面等式中便得到 $\varphi_t \circ \varphi_{-t}$=id. 因而对每个 t, 映射 φ_t 是个微分同胚, 并且 $(\varphi_t)^{-1} = \varphi_{-t}$.

现在我们给出一些单参数群的例子.

例题 5.1 设 $M = V$ 为一个向量空间, $w \in M$. 设 $\varphi : V \times \mathbb{R} \to V$ 为

$$\varphi(v, t) = v + tw.$$

容易验证 (i),(ii) 和 (iii) 都被满足 (见图 9.10).

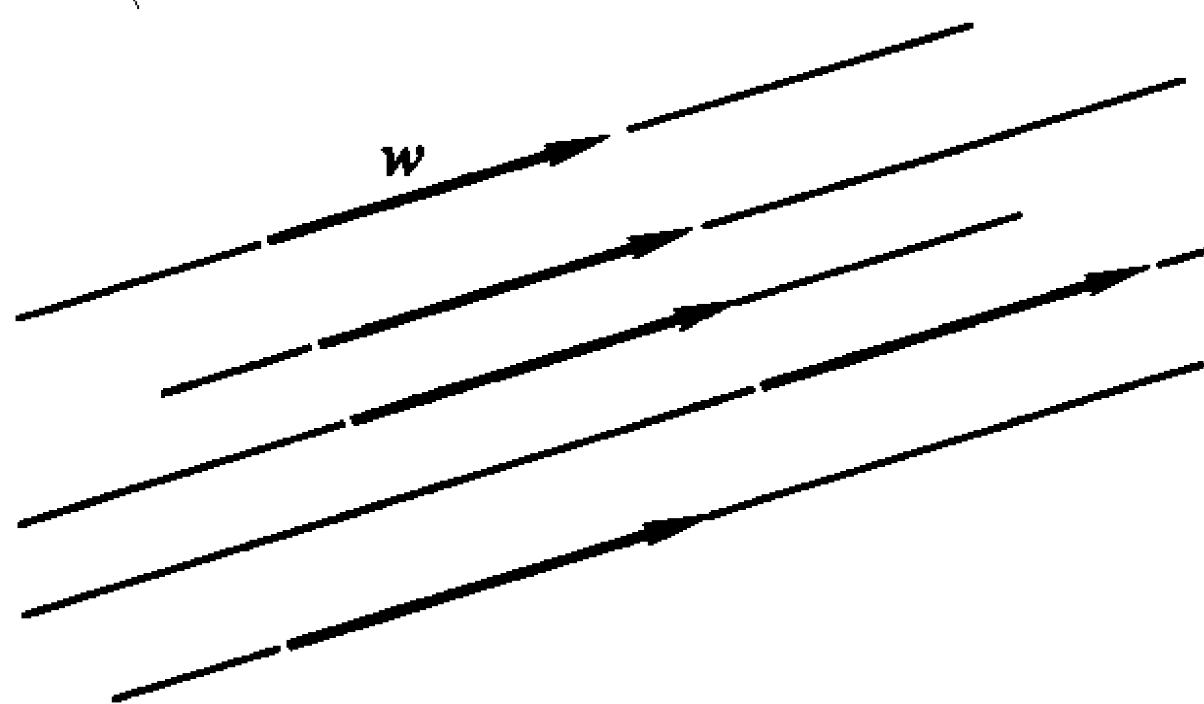


图 9.10

例题 5.2 设 $M = V$ 为有限维向量空间, $A : V \to V$ 为线性变换. 回忆线性变换 e^{tA}, 它的定义是

$$e^{tA} = 1 + tA + \frac{t^2A^2}{2!} + \frac{t^3A^3}{3!} + \cdots;$$

即, 对任意 $v \in V$,

$$e^{tA}v = \sum_{j=0}^{\infty} \frac{t^j}{j!} A^j v$$

(见图 9.11 和 9.12). 由于在任意 $\prec v, t \succ$ 的紧集合上此级数的收敛性是一致的, 所以由

$$\varphi(v, t) = e^{tA}v$$

定义的映射 $\varphi : M \times \mathbb{R} \to M$ 容易看出是可微的并同时满足 (ii) 和 (iii).

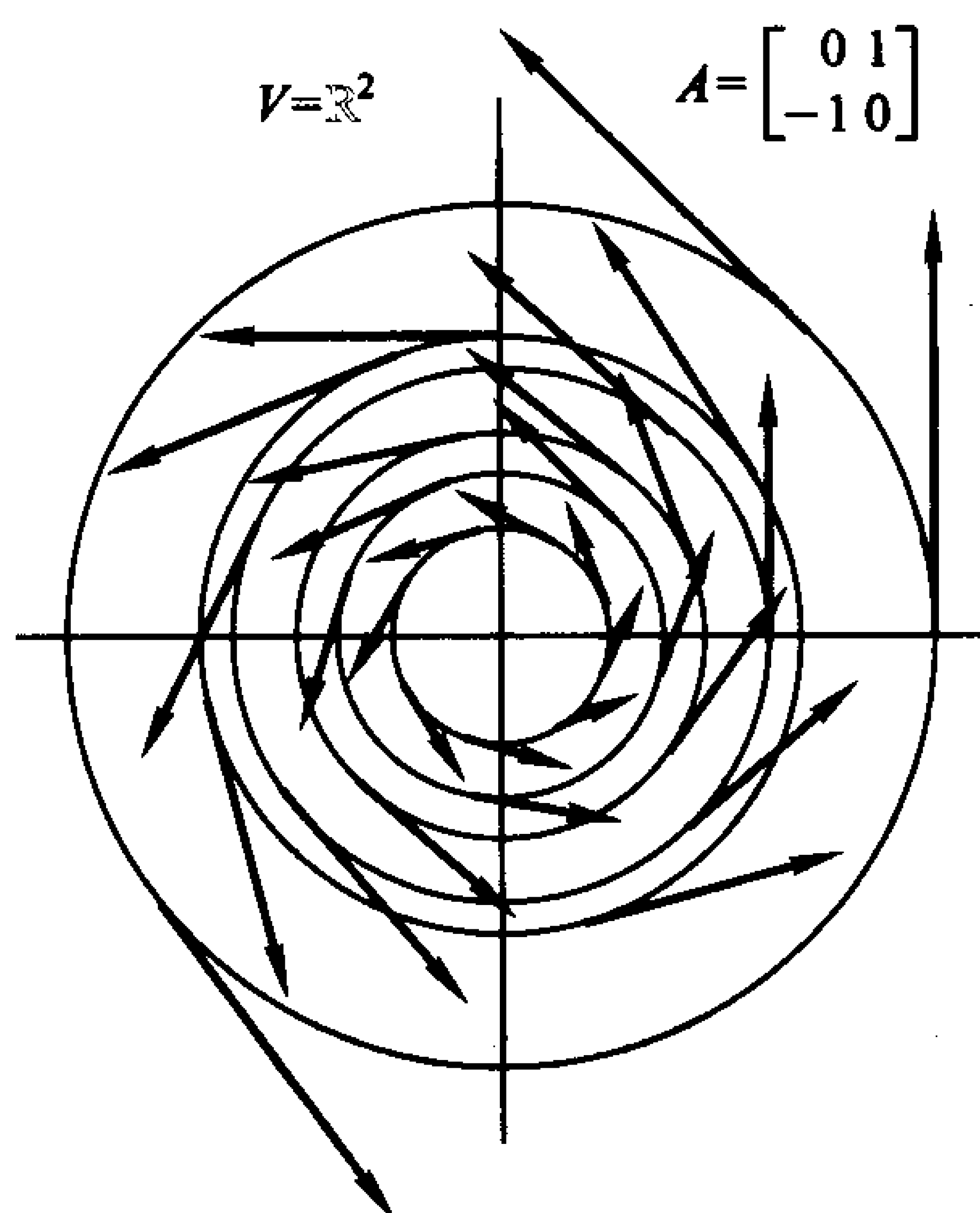


图 9.11

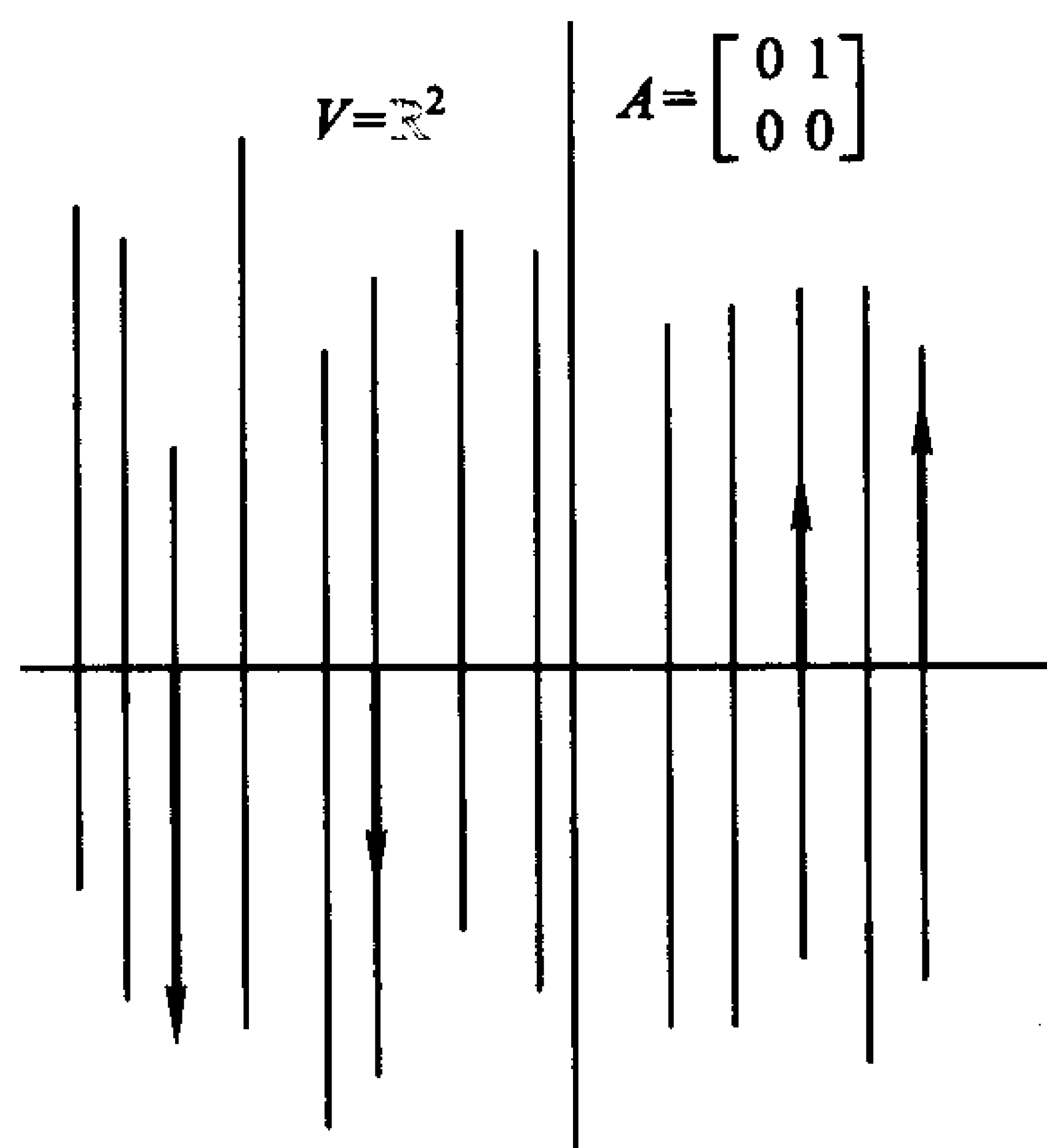


图 9.12

例题 5.3 设 M 为圆 S^1, 并设 a 为任意实数. 设 φ_t^a 由转旋角 ta 的变换组成的微分同胚. 利用总图表 $\mathfrak{a}=\{(U_1,\theta_1),(U_2,\theta_2)\}$, 映射 φ 由下面给出:

$$\begin{aligned}
\theta_1(\varphi(x,t)) &= \theta_1(x)+ta, && \text{当} x\in U_1, \theta_1(x)<2\pi-ta,\\
&= \theta_1(x)+ta-2\pi, && \text{当} x\in U_1, \theta_1(x)>2\pi-ta,\\
\theta_2(\varphi(x,t)) &= \theta_2(x)+ta, && \text{当} x\in U_2, \theta_2(x)<2\pi+\pi/2-ta,\\
&= \theta_2(x)+ta-2\pi, && \text{当} x\in U_2, \theta_2(x)>2\pi+\pi/2-ta.
\end{aligned}$$

(严格地说, 这并没有对 $\prec x,t\succ$ 的所有值全都定义 φ. 如果 $x=\prec 1,0\succ$, 并且 $ta=\dfrac{\pi}{2}$, 则 $x\notin U_1$ 且 $\varphi(x,\dfrac{\pi}{2})\notin U_2$. 这容易由引进第三个分图表予以修补.) 这容易看出 φ 为一个单参数群.

例题 5.4 设 $M=S^1\times S^1$ 为环面, a 和 b 为实数. 记 $x\in M$ 为 $x=\prec x_1,x_2\succ$, 其中 $x_i\in S^1$. 定义 $\varphi^{\prec a,b\succ}$ 为

$$\varphi^{\prec a,b\succ}(x_1,x_2,t)=\prec\varphi_t^a(x_1),\varphi_t^b(x_2)\succ,$$

其中 φ^a 和 φ^b 由例 5.3 给出. 于是 $\varphi^{\prec a,b\succ}$ 是个单参数群, 这的确是个十分有益的例题. 读者应该验证并看出本质上的行为差异来自 b/a 是有理数还是无理数.

[来自例 5.3 的例 5.4 的构造可作如下的推广. 如果 $\varphi: M\times\mathbb{R}\to M, \psi: N\times\mathbb{R}\to N$ 为单参数群, 则我们可以构造出由 $\varphi_t\times\psi_t$ 给出的 $M\times N$ 上的单参数群. 映射 $M\times N\times\mathbb{R}\to M\times N$ 将 $\prec x,y,t\succ$ 送到 $\prec\varphi_t(x),\psi_t(y)\succ$ 是可微的, 这是因为它可以被写为复合 $(\varphi\times\psi)\circ\Delta$, 其中

$$\varphi\times\psi: M\times\mathbb{R}\times N\times\mathbb{R}\to M\times N$$

而

$$\Delta: M\times N\times\mathbb{R}\to M\times\mathbb{R}\times N\times\mathbb{R}$$

为 $\Delta(x,y,t)=\prec x,t,y,t\succ$.]

在前面四个例题的每一个中, 我们都从一个"无穷小生成元"出发构造了单参数群, 就是例 5.1 中的向量 w, 例 5.2 中的线性变换 A, 例 5.3 中的实数 a, 以及例 5.4 中的偶对 $\prec a,b\succ$. 现在我们要证明, 关于流形上的任何一个单参数群, 存在一个合适的东西, 我们可以将其视为此单参数群的无穷小生成元.

设 $\varphi: M\times\mathbb{R}\to M$ 为一单参数群. 对每个 $x\in M$, 考虑映射 $\varphi_x:\mathbb{R}\to M$, 它由

$$\varphi_x(t)=\varphi(x,t)$$

给出. 由于条件 (ii) 知 $\varphi_x(0)=x$. 因此 φ_x 是通过 x 的一条曲线 (见图 9.13). 以 $X(x)$ 表示此曲线的切向量. 因此我们得到了一个映射 X, 它对每个 $x\in M$ 指定了一个向量 $X(x)\in T_x(M)$. 任意一个这样的映射, 即任何一个对每个 $x\in M$ 指定一个向量 $X(x)\in T_x(M)$ 的规则, 都被称作为一个向量场. 我们已经看到了, 每个单参数群产生了一个向量场. 我们称其为这个单参数群的无穷小生成元.

设 Y 为 M 上一个向量场,(U,α) 为 M 上一个分图表. 对每个 $x\in U$, 我们得到一个向量 $Y(x)_\alpha\in V$. 我们可以将此看作在 $\alpha(U)$ 上定义的一个 V - 值函数 Y_α: 对 $v\in\alpha(U)$,

$$Y_\alpha(v)=Y(\alpha^{-1}(v))_\alpha. \tag{5.1}$$

设 (W,β) 为第二个分图表, 并设 Y_β 为在 $\beta(W)$ 上相对应的 V - 值函数. 如果我们比较 (5.1) 和 (4.3) 式, 我们看到有

$$Y_\beta(\beta\circ\alpha^{-1}(v))=J_{\beta\circ\alpha^{-1}}(v)\circ Y_\alpha(v), \tag{5.2}$$

其中 $v\in\alpha(U\cap W)$.

等式 (5.1) 给出了一个向量场关于一个分图表的"局部表达式", 而等式 (5.2) 描绘了从一个分图表到另一个的"转移法则".

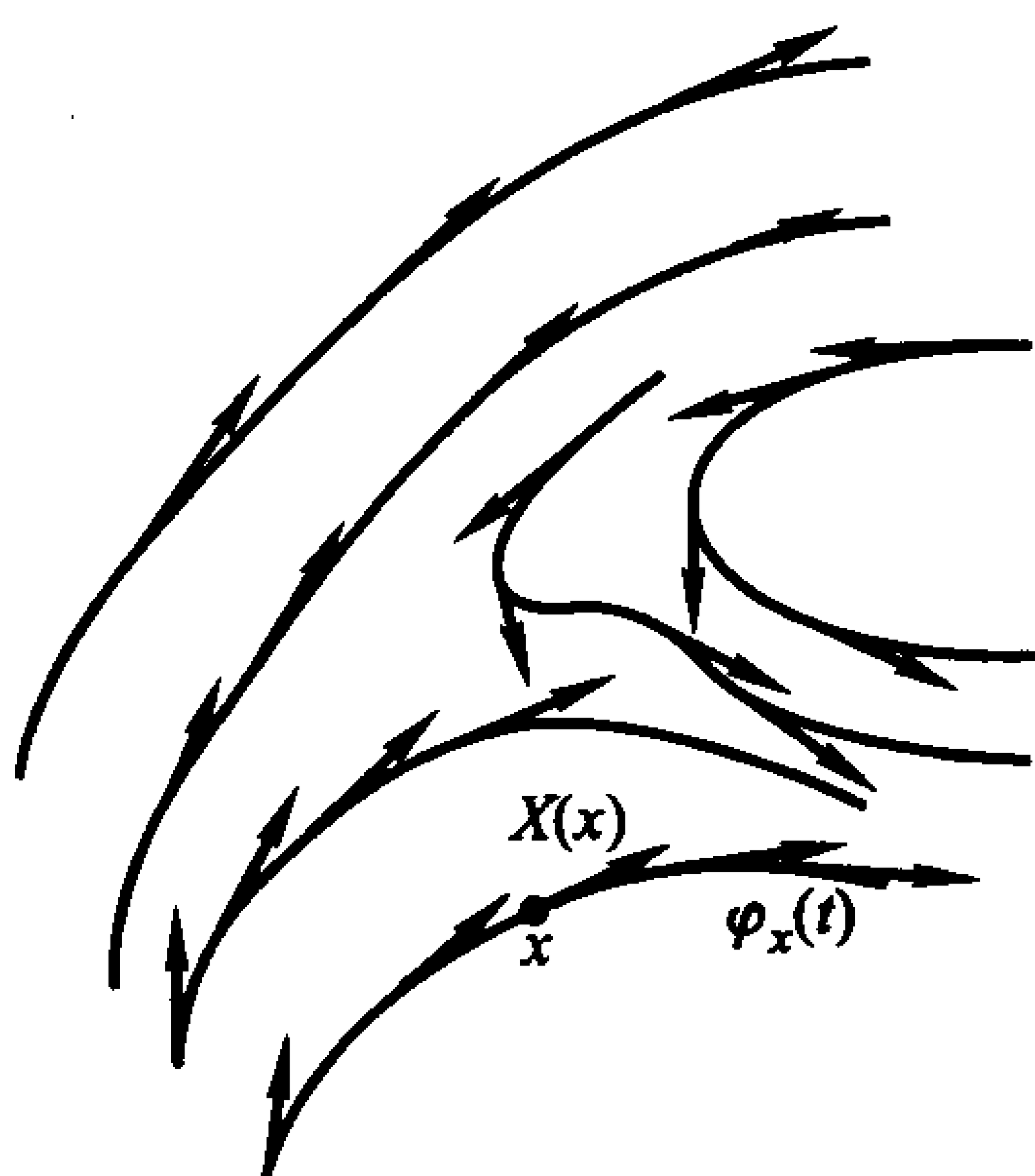


图 9.13

反过来, 设 𝔞 为 M 的一个总图表, Y_α 为对每个分图表 $(U,\alpha)\in\mathfrak{a}$ 定义在 $\alpha(U)$ 上的 V - 值函数. 假设 Y_α 满足 (5.2) 式. 于是对每个 $x\in M$, 我们可以设 $Y(x)\in T_x(M)$ 为

$$Y(x)_\alpha = Y_\alpha(\alpha(x))$$

所定义, 其中 (U,α) 为某个关于 x 的分图表. 由 (4.3) 和 (5.2) 式给出的转移法则知道此定义不依赖于 (U,α) 的选取.

又可知 $J_{\beta\circ\alpha^{-1}}$ 是 $\alpha(U\cap W)$ 上的 C^∞ 函数 (线性变换 - 值的函数). 因此, 如果 Y 是个向量场, Y_α 则是个 $\alpha(U)$ 上的 V - 值 C^∞ 函数, 函数 Y_β 则是 $\beta(U\cap V)$ 上的 C^∞ 函数. 换句话说, 它与要求函数 Y_α 为 C^∞ 类是一致的. 我们因而说, 如果对每个分图表 (U,α) 函数 Y_α 是 C^∞ 的则说 Y 是 C^∞ 向量场. 与函数和映射的情形一样, 为了验证 Y 是 C^∞ 的, 只要对属于 M 的某个总图表的所有分图表 (U,α) 验证 Y_α 是 C^∞ 的即可.

我们来验证单参数群 φ 的无穷小生成元 X 是个 C^∞ 向量场. 事实上, 如果 (U,α) 是个分图表, 则

$$X_\alpha(v) = (\alpha\circ\varphi_x)'(0),$$

其中 $\varphi_x(t)=\varphi(x,t)$. 我们可以写成 $\alpha\circ\varphi_x(t)=\Phi_\alpha(v,t)$, 其中

$$\Phi_\alpha(v,t) = \alpha\circ\varphi(\alpha^{-1}(v),t).$$

设 $U'\subset U$ 是 x 的一个邻域使得对 $y\in U'$ 和 $|t|<\varepsilon$ 有 $\varphi(y,t)\in U$. 于是 Φ_α 是 $\alpha(U')\times I\to\alpha(U)$ 的一个可微函数, 其中 $I=\{t:|t|<\varepsilon\}$. 用这个表示方法, 我

们可写成

$$X_\alpha(v) = \frac{\partial \Phi}{\partial t}(v, 0). \tag{5.3}$$

这表明 X 是个 C^∞ 向量场.

如果在例 1 的情形中计算 (5.3) 式的值, 我们得到 $\Phi_{\rm id}(v,t) = v + tw$, 故 $X_{\rm id} = w$. 在例 2 的情形有 $X_{\rm id}(v) = Av$.

对向量场可以进行各种代数运算. M 上所有向量场的集合构成了一个向量空间, 其方式是显然的. 如果 X 和 Y 为 C^∞ 向量场, 则 $aX + bY$ 也是 (a, b 为常数), 其中

$$(aX + bY)(x) = aX(x) + bY(x),\ x \in M.$$

相似地, 我们可以对向量场乘以一个函数. 如果 f 是个函数, X 是个向量场. 定义 fX 为

$$(fX)(x) = f(x)X(x),\ x \in M.$$

容易看出如果 f 和 X 是可微的, 则 fX 也是. 也容易验证对此乘法的各种结合律也成立.

我们已经知道任何一个单参数群定义了一个光滑向量场. 让我们检验相反的情形. 是否任一个 C^∞ 向量场都定义了一个单参数群? 所说的这个问题的答案为"否".

事实上, 设 $X = \dfrac{\partial}{\partial x^1}$ 为 $\mathbb{R}^n$ 中对应 x^1 方向平移的向量场. 设 $M = \mathbb{R}^2 - C$, 其中 C 为 $\mathbb{R}^2$ 中某个非空闭集. 那么, 如果 p 是 M 中位于平行于 x^1 轴且与 C 相交的直线上的一点 (图 9.14), 则 $\varphi_t(p)$ 不是对每个 t 都有定义 (它将不在 M 中).

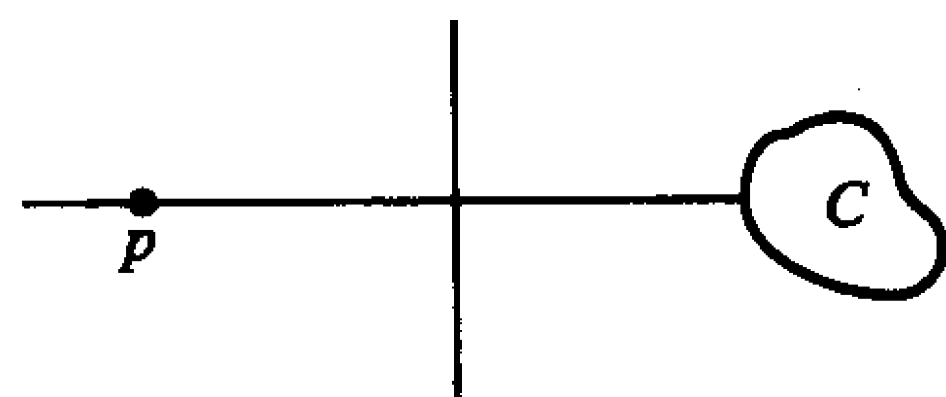


图 9.14

读者可能不喜欢这种"丢掉某些点"的 M, 并说这就是 X 不能生成单参数群的原因. 但我们可以在 $\mathbb{R}^2$ 自身上构造一个反例. 事实上, 如果考虑 $\mathbb{R}^2$ 上向量场 X, 它由

$$X_{\rm id}(x^1, x^2) = (1, -(x^2)^2)$$

给出. 那么, (5.3) 式表明如果 φ 有定义, 则满足

$$\frac{d\Phi}{dt}(x,t) = \frac{d\Phi}{dt}(\Phi(x,t), 0) = X(\Phi(x,t)),$$

其中 $\Phi = \Phi_{\mathrm{id}}$. 如果设 $y^i(t,x) = x^i \circ \Phi(x,t)$, 则

$$\frac{dy^1}{dt} = 1, y^1(0) = x^1,$$
$$\frac{dy^2}{dt} = -(y^2)^2, y^2(0) = x^2.$$

如果 $x^2 \neq 0$, 则第二个方程的惟一解由

$$y^2(t) = \frac{1}{t + 1/x^2}$$

给出, 它不是对所有的 t 都有定义. 当然, 麻烦是由于我们只有一个微分方程的局部存在性定理而引起的.

因此, 我们必须放弃对 φ 要定义在整个 $M \times \mathbb{R}$ 上的要求.

定义 5.1 M 上一个流是从开集 $U \subset M \times \mathbb{R}^1$ 到 M 的一个映射 φ, 使得

(i) $M \times \{0\} \subset U$;

(ii) φ 可微;

(iii) $\varphi(x,0) = x$;

(iv) $\varphi(\varphi(x,s),t) = \varphi(x,s+t)$, 只要等式两端的映射都存在, 则等式成立.

对固定的 $x, \varphi_x(t) = \varphi(x,t)$ 对充分小的 t 有定义, 故 φ 产生了一个像从前那样的向量场 X. 我们称 X 为流 φ 的无穷小生成元.

像先前的例子表明的, 会发生没有 $t \neq 0$ 使得 $\varphi(x,t)$ 对所有的 x 有定义, 也会发生没有 x 使 $\varphi(x,t)$ 对所有的 t 都有定义.

命题 5.1 设 X 为 M 上一个光滑向量场. 则存在 $M \times \{0\}$ 在 $M \times \mathbb{R}$ 中一个邻域 U 及一个流 $\varphi : U \to M$, 它以 X 为其无穷小生成元.

证明 首先对任意 $x \in M$ 构造曲线 $\varphi_x(t)$, 而后验证 $\prec x, t \succ \mapsto \varphi(x,t)$ 的确是个流. □

设 x 是 M 中一点, (U,α) 是 x 的一个分图表. 则 X_α 给了我们一个在 $\alpha(U)$ 中的常微分方程, 即

$$\frac{dv}{dt} = X_\alpha(v), v \in \alpha(U).$$

由常微分方程的基本存在定理, 存在一个 $\varepsilon > 0$ 及包含 $\alpha(x)$ 的开集 O 和一个映射

$$\Phi_\alpha : O \times \{t : |t| < \varepsilon\} \to \alpha(U)$$

使得

$$\Phi_\alpha 为 C^\infty, \Phi_\alpha(v,0)=v,$$

以及

$$\frac{d\Phi_\alpha(v,t)}{dt}=X_\alpha(\Phi_\alpha(v,t)).$$

这里开集 O 和 ε 的选取依赖于 $\alpha(x)$ 和 $\alpha(U)$. 此基本定理的惟一性部分断言 Φ_α 在定义域内是惟一确定的, 即如果 Φ_v 是任何对 $|t|<\varepsilon'$ 有定义的曲线且 $\Phi_v(0)=v$, 以及

$$\frac{d\Phi_v(t)}{dt}=X_\alpha(\Phi_v(t)), \tag{5.4}$$

则 $\Phi_v(t)=\Phi_\alpha(v,t)$.

这表明

$$\Phi_\alpha(v,t+s)=\Phi_\alpha(\Phi_\alpha(v,s),t)$$

只要在两端都有定义就成立.(正好在方程中固定 s 时成立.)

考虑曲线 $\varphi_x^\alpha(\cdot)$, 它由

$$\varphi_x^\alpha(t)=\alpha^{-1}(\Phi_\alpha(\alpha(x),t)) \tag{5.5}$$

定义 5.2 当 $|t|<\varepsilon$ 它有定义并且连续, 事实上是这个区间到 M 内的可微映射. 进一步, 如果我们记 $\psi=\varphi_x^\alpha(\cdot)$, 则 (5.4) 式断言了曲线 $\psi(t+\cdot)$ 的切向量是 $X(\psi(t))$, 即此向量场在点 $\psi(t)$ 的值. 我们将此条件写为

$$\psi'(t)=X(\psi(t)). \tag{5.6}$$

方程 (5.6) 就是我们想写出对应于向量场 X 的 M 上"一阶微分方程"的方法. 一条满足 (5.6) 式的可微曲线 ψ 被称做 X 的积分曲线. 我们现在要正式叙述微分方程的惟一性定理的流形形式:

引理 5.1 设 $\psi_1:I\to M,\psi_2:I\to M$ 为定义在同一区间 I 上的 X 的两条积分曲线. 如果 $\psi_1(s)=\psi_2(s)$ 对某个点 $s\in I$ 成立, 则 $\psi_1=\psi_2$, 即对所有 $t\in I$ 有 $\psi_1(t)=\psi_2(t)$.

证明 我们想证明 $\psi_1(t)\neq\psi_2(t)$ 的集合为空. 设

$$A=\{t:t\geqslant s 且 \psi_1(t)\neq\psi_2(t)\}.$$

我们要让 A 为空集, 并且相似地, 证明集 $B=\{t: t \leqslant s$ 且 $\psi_1(t) \neq \psi_2(t)\}$ 为空集. 假设 A 非空, 并设

$$t_+ = \text{glb}A = \text{glb}\{t : t \geqslant s\text{且}\psi_1(t) \neq \psi_2(t)\}.$$

我们要以下面的方式推导出矛盾:

(i) 应用微分方程的惟一性定理证明 $\psi_1(t_+) \neq \psi_2(t_+)$, 以及

(ii) 应用流形的豪斯多夫性质证明 $\psi_1(t_+) = \psi_2(t_+)$.

证明的细节: (i) 假设 $\psi_1(t_+) = \psi_2(t_+) = y \in M$. 我们可以找到 y 的一个坐标分图表 (β, W), 于是 $\beta \circ \psi_1$ 和 $\beta \circ \psi_2$ 是同一个一阶常微分方程组的解, 它们在 $t = t_+$ 取相同的值 $\beta(y)$. 因而由微分方程解的惟一性, $\beta \circ \psi_1$ 和 $\beta \circ \psi_2$ 必在 t_+ 的某个区间内相等, 从而对此区间的所有 t 有 $\psi_1(t) = \psi_2(t)$. 因为由 t_+ 的上确界性质, 必定有点任意地靠近 t_+, 并且在这些点有 $\psi_1(t) \neq \psi_2(t)$, 故这是不可能的. 这就证明了 (i). 现在假设 $\psi_1(t_+) \neq \psi_2(t_+)$. 我们可以找到 $\psi_1(t_+)$ 的邻域 U_1 和 $\psi_2(t_+)$ 的邻域 U_2, 使得 $U_1 \cap U_2 = \phi$. 然而 ψ_1 的连续性意味着当 t 十分靠近 t_+ 时有 $\psi_1(t) \in U_1$ 和 $\psi_2(t) \in U_2$, 从而对 t_+ 的某个区间中的 t 有 $\psi_1(t) \neq \psi_2(t)$, 这再一次与 t_+ 的上确界性质相抵触因而证明了 (ii). 将 glb 换作 lub 作同样的论证便证明了 B 是空集, 故证明了引理 5.1. 上面的论证是典型的"连通性论证法". 我们证明了使 $\psi_1(t) = \psi_2(t)$ 的集合既开又闭, 从而必是整个区间 I.

引理 5.1 表明 (5.5) 式定义了在时间 $t=0$ 时通过 x 点的 X 的一条解曲线, 并且在 O 的任何一个公共的定义区间上与分图表的选取无关. 换句话说, 以

$$\varphi_x(t) = \alpha^{-1}(\Phi(\alpha(s), t))$$

定义曲线 $\varphi_x(\cdot)$, 它对 $|t| < \varepsilon, \varphi_x(t)$ 有定义. 可惜 ε 不仅依赖于 x 而且也依赖于图表的选取. 我们利用引理 5.1 尽可能地扩张 $\varphi_x(\cdot)$ 的定义, 极像我们在第六章中对向量空间上常微分方程所做过的那样: 对任意满足 $|s| < \varepsilon$ 的 s, 设 $y = \varphi_x(s)$, 并对 $|t| < \varepsilon'$ 得到一条在其上有定义的曲线 $\varphi_y(\cdot)$. 由引理 5.1 得到如果 $|s+t| < \varepsilon$ 有

$$\varphi_y(t) = \varphi_x(s+t). \tag{5.7}$$

可能出现 $|s| + \varepsilon' > \varepsilon$. 于是会存在一个 t, 满足 $|t| < \varepsilon'$ 而 $|s+t| > \varepsilon$. 那么, (5.7) 式的右端没有定义, 但左端却有. 于是我们取 (5.7) 式为 $\varphi_x(s+t)$ 的定义, 从而拓展了 $\varphi_x(\cdot)$ 的定义域. 然后我们继续进行: 设 I_x^+ 表示所有 $s > 0$ 并且存在有限的实数序列 $s_0 = 0 < s_1 < \cdots < s_k = s$ 及点 $x_0, \cdots, x_{k-1} \in M$, 其中

$x_0 = x, s_1$ 在 $\varphi_x(\cdot)$ 的定义域中, $x_2 = \varphi_x(s_1)$, 等等, 归纳地, s_{i+1} 在 $\varphi_{x_i}(\cdot)$ 的定义域中且 $x_{i+1} = \varphi_{x_i}(s_{i+1})$, 这样的 s 的集合.

如果 $s \in I_x^+$, 故对 $0 < s' < s$ 的 s' 也如此, 于是对充分小的正数 $\eta, s+\eta$ 也属于 I_x^+. 因此 I_x^+ 是个区间, 其半开于右端. 反复运用 (5.4) 式, 我们对 $s \in I_x^+$ 定义了 $\varphi_x(s)$. 以相似的办法构造 I_x^- 并令 $I_x = I_x^+ \cup I_x^-$. 于是 $\varphi_x(s)$ 对 $s \in I_x$ 有定义, 并且 I 是我们构造所定义的 φ_x 的最大区间. 对每个 $x \in M$, 我们得到一个开区间 I_x, 使在其中曲线 $\varphi_x(\cdot)$ 有定义.

设 $U = U_{x\in M}\{x\} \times I_x$. 于是 U 是个 $M \times I$ 的开子集. 要证实这点, 我们设 $(\overline{x}, \overline{s}) \in U$. 我们必须证明存在 $\overline{x}$ 的一个邻域 W 和一个 $\varepsilon > 0$, 使得对所有 $|s - \overline{s}| < \varepsilon$ 和 $x \in W$ 有 $s \in I_x$. 由定义, 存在有限的点序列 $\overline{x} = \overline{x}_0, \overline{x}_1, \cdots, \overline{x}_k$ 和分图表 $(U_1, \alpha_1), \cdots, (U_k, \alpha_k)$, 满足 $x_{i-1} \in U_i, x_i \in U_i$ 使得

$$\alpha_i(\overline{x}_i) = \Phi_{\alpha_i}(\alpha_i(\overline{x}_{i-1}), t_i),$$

其中 $t_1 + \cdots + t_k = s$. 由 Φ_α 的连续性现在能清楚知道, 如果我们选取 x_0 使得 $\alpha_1(x_0)$ 离 $\alpha_1(\overline{x}_0)$ 足够近, 则由

$$\alpha_i(x_i) = \Phi_{\alpha_i}(\alpha_i(x_{i-1}), t_i)$$

归纳定义的点 x_i 是有明确定义的.[那就是说,$\alpha_j(x_{j-1})$ 在 $\Phi_{\alpha_i}(\cdot, t_i)$ 的定义域中.] 这意味着对所有这样的点 $x_0, \overline{s} \in I_x$. 相同的论证表明对 η 充分小以及 x 充分靠近 $\overline{x}$ 有 $\overline{s} + \eta \in I_x$. 这就证明了 U 是开的.

现对 $(x,t) \in U$ 令

$$\varphi(x,t) = \varphi_x(t),$$

它定义了 $\varphi \circ \varphi$ 在 $M \times \{0\}$ 附近可微性由 φ 作为一个常微分方程的解而得到 (借助于一个分图表). 基本存在性定理保证了可微性. 在点 $(\overline{x}, \overline{t})$ 附近我们可以写成

$$\varphi(x,t) = \varphi(\varphi(\cdots(\varphi(x,t_1),t_2)\cdots)t_k), t = t_1 + t_2 + \cdots + t_k,$$

故由它是可微映射的复合得到 φ 是可微的. □

9.6 李导数

设 φ 为流形 M 上的单参数群, f 为 M 上的可微函数. 于是对每个 t, 函数 $\varphi_t^*[f]$ 是可微的, 并且对 $t \neq 0$ 我们可以构造函数

$$\frac{\varphi_t^*[f] - f}{t}. \tag{6.1}$$

我们断言, 当 $t \to 0$ 时这个表达式的极限存在. 事实上, 对任意 $x \in M, \varphi_t^*[f](x) = f \circ \varphi_t(x) = f \circ \varphi_x(t)$, 因此

$$\begin{aligned}\lim_{t\to 0}\frac{\varphi_t^*[f]-f}{t}(x) &= \lim_{t\to 0}\frac{f\circ\varphi_x(t)-f\circ\varphi_x(0)}{t}\\ &= D_{\varphi_x}f = X(x)f.\end{aligned} \tag{6.2}$$

这里的 $X(x)$ 是在 x 的一个切向量, 我们所使用的是在第 4 节引进的记号. 我们称 (6.1) 式的极限为 f 关于单参数群 φ 的导数, 记为 $D_X f$. 更一般地, 对任意一个光滑向量场 X 和可微函数 f, 我们定义 $D_X f$ 为

$$D_X f(x) = X(x)f, \text{对所有的} x \in M, \tag{6.3}$$

并称其为 f 关于 X 的李导数. 用 X 生成的流来表示, 我们可以在任一 $x \in M$ 附近把 $D_X f$ 表示为 (6.1) 式的极限, 一般说来, (6.1) 式在这里只对 x 的充分小邻域和充分小的 $|t|$ 有定义.

我们的记号推广了在第三章中对方向导数的记号. 事实上, 如果 M 是 V 中一个开集, X 是例 1 中那个 "常值向量场"

$$X_{\mathrm{id}} = w \in V,$$

则

$$(D_X f)_{\mathrm{id}} = D_w f_{\mathrm{id}},$$

其中 D_w 是关于 w 的方向导数.

注意, $D_X f$ 对 X 是线性的; 就是说,

$$D_{aX+bY}f = aD_X f + bD_Y f,$$

其中 X 和 Y 是向量场. a 和 b 是常数.

设 ψ 是 M_1 到 M_2 上的微分同胚, X 为 M_2 上的一个向量场. 我们定义在 M_1 上的一个 "拉回" 向量场 $\psi^*[X]$ 为, 对所有 $x \in M$,

$$\psi^*[X](x) = \psi_*^{-1}X(\psi(x)). \tag{6.4}$$

注意, 要使 (6.4) 式有意义, ψ 必须是微分同胚, 这是因为定义涉及到了 ψ^{-1}. 这与函数的 "拉回" 成了对照, 它对任何可微映射都有意义. 等式 (6.4) 的确定义了一个向量场. 理由是

$$\psi_*^{-1} = \psi_{*x}^{-1} : T_{\psi(x)}(M_2) \to T_x(M_1), \text{且} X(\psi(x)) \in T_{\psi(x)}(M_2).$$

让我们验证: 如果 X 是个光滑向量场则 $\psi^*[X]$ 也是. 其大意是, 设 $\mathfrak{a}_1$ 和 $\mathfrak{a}_2$ 是 M_1 和 M_2 上相容的总图表, 并设 $(U,\alpha) \in \mathfrak{a}_1, (W,\beta) \in \mathfrak{a}_2$ 为相容的分图表. 那么 (6.4) 式说

$$\psi^*[X]_\alpha(v) = J_{\alpha\circ\psi^{-1}\circ\beta^{-1}}(\beta\circ\psi\circ\alpha^{-1}(v)) \cdot X_\beta(\beta\circ\psi\circ\alpha^{-1}(v))$$

是 v 的可微函数, 其中 $v \in \alpha(U)$. 由链规则, 因为

$$J_{\alpha\circ\psi^{-1}\circ\beta^{-1}}(\beta\circ\psi\circ\alpha^{-1}(v)) \cdot J_{\beta\circ\psi\circ\alpha^{-1}}(v) = 1,$$

所以我们可将上面最后一个表达式重写为更简单的

$$\psi^*[X]_\alpha(v) = (J_{\beta\circ\psi\circ\alpha^{-1}}(v))^{-1} X_\beta(\beta\circ\psi\circ\alpha^{-1}(v)), \tag{6.5}$$

其中 $v \in \alpha(U)$.

因此 $\psi^*[X]_\alpha$ 是光滑的 $\mathrm{Hom}(V_2, V_1)$ - 值函数和一个光滑的 V_2 - 值函数的乘积, 它证明了 $\psi^*[X]$ 是个光滑向量场.

习题

设 φ 为在 M_2 上 X 生成的流. 证明由 $\psi^*[X]$ 所生成的流由下式给出:

$$\prec x, t \succ \mapsto \psi^{-1}\varphi(\psi(x), t). \tag{6.6}$$

如果 φ 是个单参数群, 则可将 (6.6) 式写为

$$\prec x, t \succ \mapsto \psi^{-1} \circ \varphi_t \circ \psi(x). \tag{6.6$'$}$$

容易验证, 如果 $\psi_1 : M_1 \to M_2$ 和 $\psi_2 : M_2 \to M_3$ 为微分同胚, Y 是 M_3 上的向量场. 则

$$(\psi_2 \circ \psi_1)^* Y = \psi_1^* \psi_2^* Y.$$

如果 f 是 M_2 上的可微函数, 则

$$D_{\psi^*[X]}(\psi^*[f]) = \psi^*(D_X f). \tag{6.7}$$

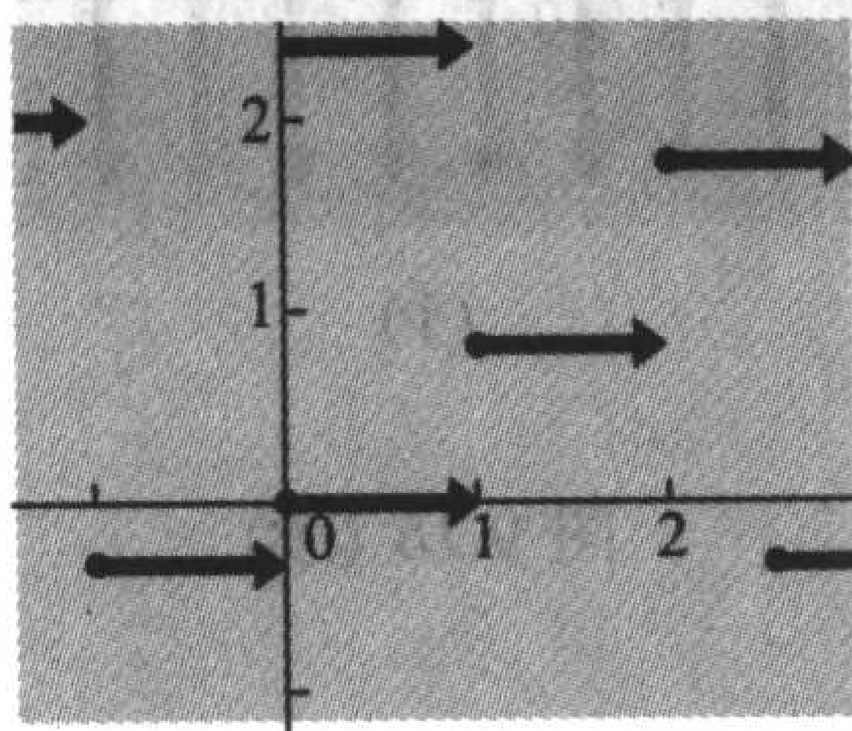
事实上, 由 (6.3) 和 (4.6) 式, 对 $x \in M_1$ 我们有

$$\begin{aligned}
D_{\psi^*[X]}\psi^*[f](x) &= \psi^*[X](x)\psi^*[f] && \text{由(6.3)}\\
&= \psi_*^{-1}X(\psi(x))\psi^*[f] && \text{由(6.4)}\\
&= (\psi_*\psi_*^{-1}X(\psi(x)))f && \text{由(4.6)}\\
&= X(\psi(x))f\\
&= D_X f)(\psi(x)) = \psi^*(D_X f)(x).
\end{aligned}$$

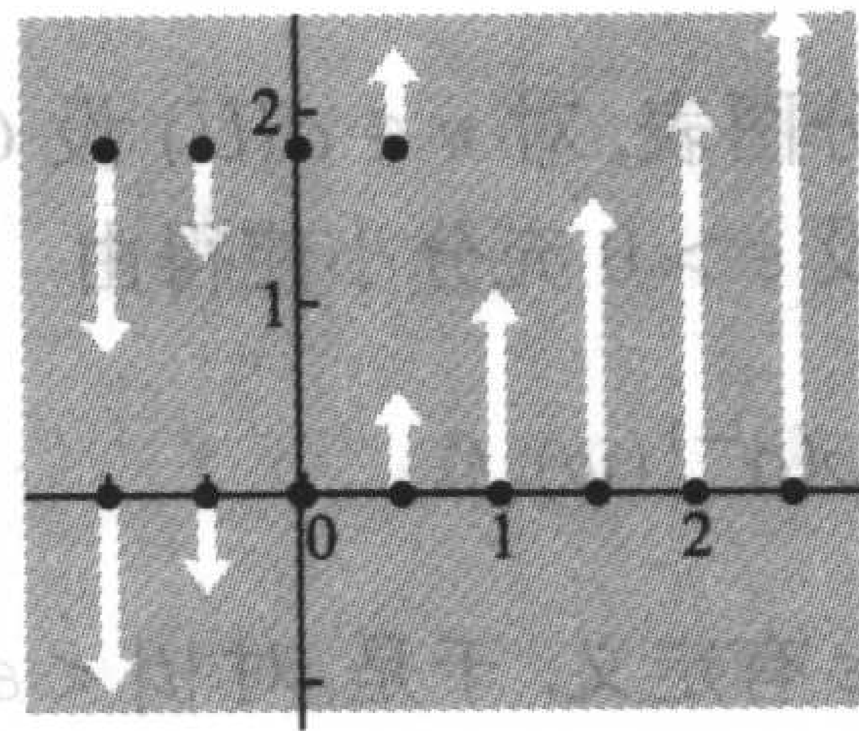
设 φ 为 M 上的单参数群, 其无穷小生成元为 X, 又设 Y 是 M 上另一个光滑向量场. 对 $t \neq 0$ 我们可以构造向量场

$$\frac{\varphi_t^*[Y] - Y}{t} \tag{6.8}$$

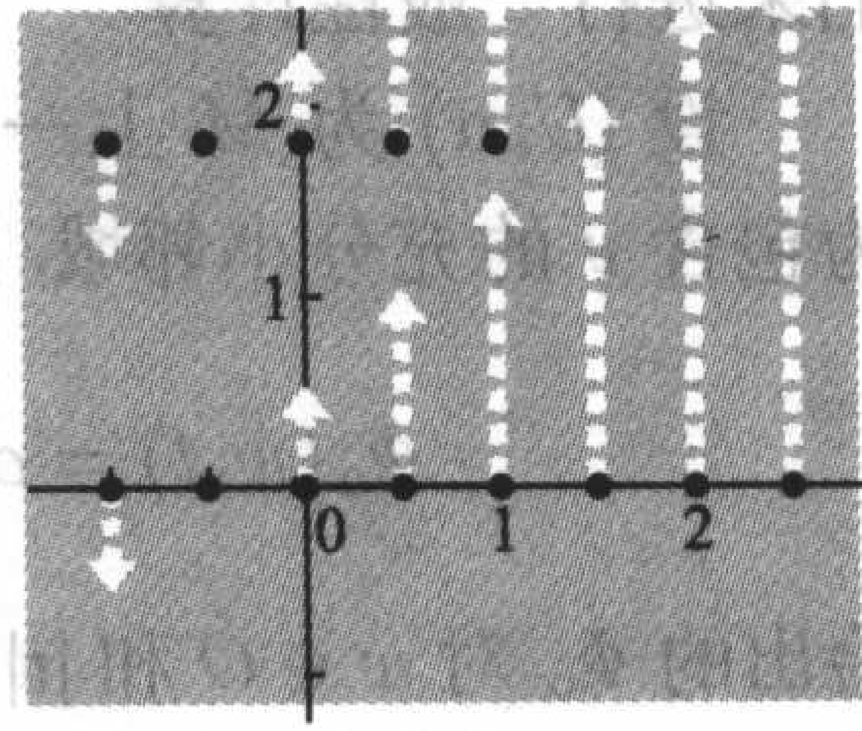
并研讨当 $t \to 0$ 时它的极限, 我们称此极限为 $D_X Y$. 在图 9.15 中我们展示了对在笛卡儿平面 $\mathbb{R}^2$ 上两个非常简单的向量场的 $D_Y X$ 和 $D_X Y$ 的计算. 向量场 X 是常值向量场 $X_{\text{id}} = \delta_1$, 故 $X_f = \partial f/\partial x$, 这是用了平面上的笛卡儿坐标表示的. 相对应的流由 $\varphi_t(x, y) = \prec x + t, y \succ$. 如果我们将在平面上每点的切平面等同于此平面自身则 φ_{t*}=id. 于是 $Y \to \varphi_t^* Y$ 由把向量场 Y 向左“移动” t 单位组成. 在这里我们已取了 $Y = x\delta_2$, 故 $Yf = x(\partial f/\partial y)$. 在图 9.15(c) 中我们画出了 $\varphi_t^* Y$, 并在图 9.15(d) 中将 Y 和 $\varphi_t^* Y$ 叠置在一起. 图 9.15(e) 代表了 $\varphi_t^* Y - Y$, 而图 9.15(f) 是 $(1/t)\{\varphi_t^* Y - Y\}$, 由于这个表达式与 t 无关, 故它与其极限 $D_X Y$ 重合. 由 Y 生成的单参数群是 $\varphi_t, \varphi_t(x, y) = \prec x, y + tx \succ$. 在任一点 $p \in \mathbb{R}^2$ 我们现在有 $\psi_{t*}\delta_1 = \delta_1 + t\delta_2$, 故 $\psi_t^* X = \psi_{-t*}X(\psi(x)) = \delta_1 - t\delta_2$.



(a)



(b)



(c)

图 9.15

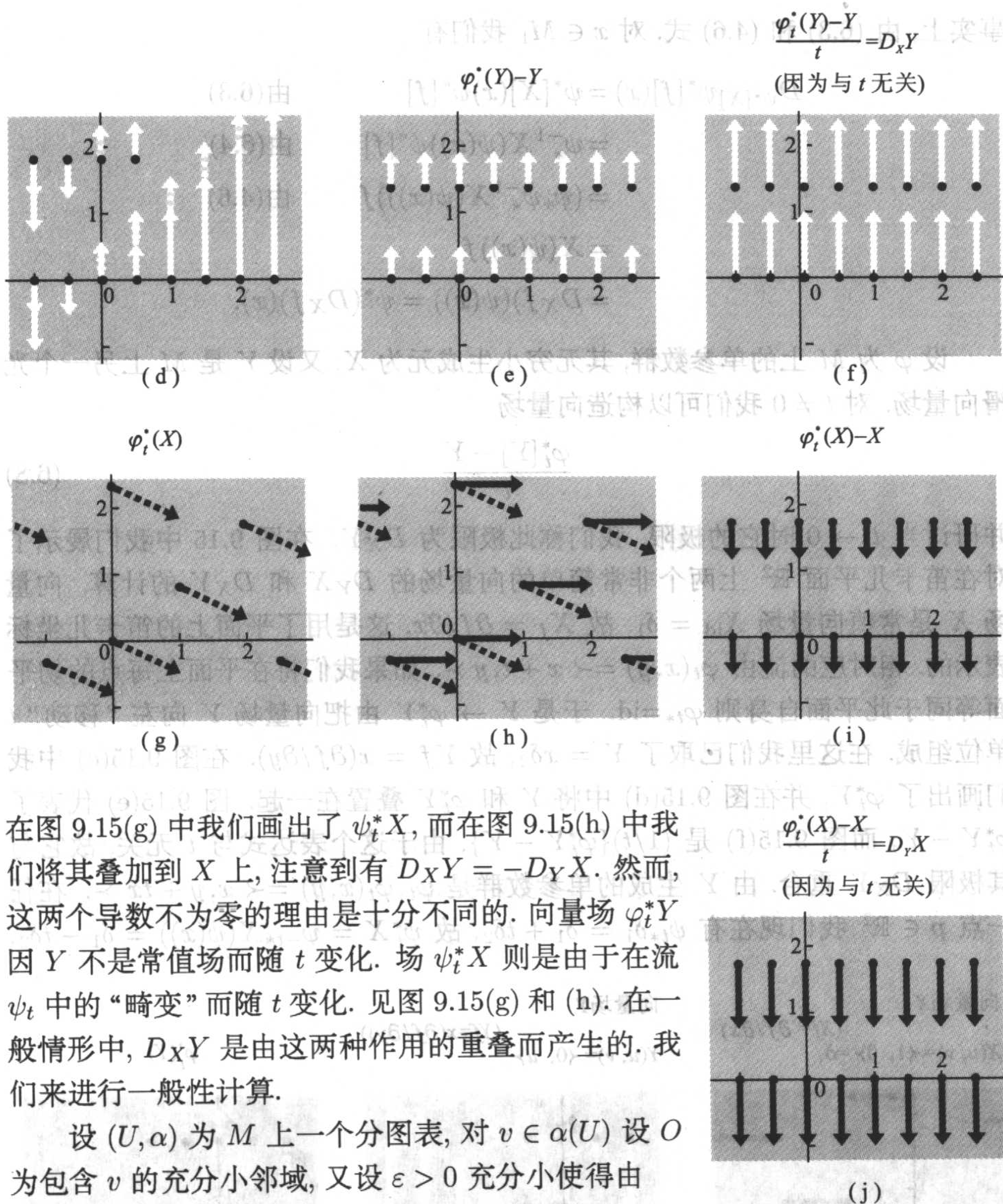


图 9.15 (续)

在图 9.15(g) 中我们画出了 ψ_t^*X, 而在图 9.15(h) 中我们将其叠加到 X 上, 注意到有 $D_XY=-D_YX$. 然而, 这两个导数不为零的理由是十分不同的. 向量场 φ_t^*Y 因 Y 不是常值场而随 t 变化. 场 ψ_t^*X 则是由于在流 ψ_t 中的"畸变"而随 t 变化. 见图 9.15(g) 和 (h). 在一般情形中, D_XY 是由这两种作用的重叠而产生的. 我们来进行一般性计算.

设 (U,α) 为 M 上一个分图表, 对 $v\in\alpha(U)$ 设 O 为包含 v 的充分小邻域, 又设 $\varepsilon>0$ 充分小使得由

$$\Phi_\alpha(w,t)=\alpha\circ\varphi(\alpha^{-1}(w),t)$$

给出的 Φ_α 对 $w\in O$ 和 $|t|<\varepsilon$ 有定义. 于是, 对 $|t|<\varepsilon$, 等式 (6.5) 蕴涵了

$$\varphi_t^*[Y]_\alpha(v)=(J_{\Phi_{\alpha(v,t)}}(v))^{-1}Y_\alpha(\Phi_\alpha(v,t)). \tag{6.9}$$

这个等式的右端具有形式 $A_t^{-1}z_t$，其中 A_t 和 z_t 是 t 的可微函数，且 $A_0 = I$. 因此它关于 t 的导数存在，并有

$$\begin{aligned}\left.\frac{d(A_t^{-1}z_t)}{dt}\right|_{t=0} &= \lim_{t\to 0}\frac{A_t^{-1}z_t - z_0}{t} = \lim_{t\to 0} A_t\frac{(A_t^{-1}z_t - z_0)}{t}\\ &= \lim_{t\to 0}\frac{z_t - A_t z_0}{t} = \lim_{t\to 0}\left(\frac{z_t - z_0}{t} - \frac{A_t z_0 - z_0}{t}\right)\\ &= z_0' - A_0' z_0.\end{aligned}$$

在 (6.9) 式中，$z_t = Y_\alpha(\Phi_\alpha(v,t))$，故

$$z_0' = dY_\alpha\left(\frac{\partial\Phi_\alpha}{\partial t}(v,0)\right) = dY_\alpha(X_\alpha(v)).$$

在这里，Y_α 是个 V - 值函数，所以 dY_α 是它在点 $\Phi_\alpha(v,0)$ 的微分. 从而 $dY_\alpha(X_\alpha(v))$ 是此微分在 $X_\alpha(v)$ 的值. 变换 $A_t = J_{\Phi_\alpha(v,t)} = d(\Phi_\alpha)_{(v,t)}$，所以

$$\left.\frac{dA_t}{dt}\right|_{t=0} = \left.\frac{\partial d\Phi_\alpha}{\partial t}\right|_{t=0} = d\frac{\partial\Phi_\alpha}{\partial t} = d(X_\alpha)_v.$$

因此 (6.9) 式在 $t=0$ 的导数可以写为

$$d(Y_\alpha)_v(X_\alpha(v)) - d(X_\alpha)_v(Y_\alpha(v)) = D_{X_\alpha(v)}Y_\alpha - D_{Y_\alpha(v)}X_\alpha.$$

我们因而证明了 (6.8) 式中的极限存在. 如果我们以 D_XY 表示它，我们则可写成

$$(D_XY)_\alpha(v) = D_{X_\alpha(v)}Y_\alpha - D_{Y_\alpha(v)}X_\alpha. \tag{6.10}$$

像前面那样，我们用 (6.10) 式作为对任意两个向量场 X 和 Y 的 D_XY 的定义. 另外这又表示了 Y 关于由 X 生成的流的导数，就是说 (6.9) 式的极限，其中 (6.8) 式现在仅为局部定义的.

从 (6.10) 式我们得到了令人吃惊的结果，即 $D_XY = -D_YX$. 由于这个原因，引进一个能更清楚表达此反对称性的记号是方便的，我们记

$$D_XY = [X,Y],$$

此表达式的右端被称做 X 和 Y 的**李括号**. 我们有

$$[X,Y] = -[Y,X]. \tag{6.11}$$

让我们来计算第 5 节开头所列出的一些例题的李括号，设 $M = \mathbb{R}^n$.

例题 6.1 如果 $X_{\mathrm{id}} = w_1, Y_{\mathrm{id}} = w_2$ 为"常值"向量场. 则 (6.10) 式表明 $[X,Y]=0$.

例题 6.2 设 $X_{\mathrm{id}}(v) = Av$, 其中 A 是个线性变换, 再设 $Y_{\mathrm{id}} = w$. 于是 (6.10) 式说

$$[X,Y]_{\mathrm{id}}(v) = -Aw,$$

这是因为线性函数 Av 的关于 w 的方向导数是 Aw.

例题 6.3 设 $X_{\mathrm{id}}(v) = Av, Y_{\mathrm{id}}(v) = Bv$, 其中 A, B 是线性变换. 于是由 (6.10) 式有

$$[X,Y]_{\mathrm{id}}(v) = BAv - ABv = (BA - AB)v \tag{6.12}$$

因此这时的 $[X,Y]$ 又由线性变换得到, 即 $BA-AB$. 此时在 A 和 B 中的反对称性相当清楚.

我们现在回到一般情形. 设 φ 为 M 上的单参数群, Y 为 M 上的光滑向量场, f 为 M 上的可微函数. 按照 (6.7) 式,

$$D_{\varphi_t^*[Y]}(\varphi_t^*[f]) = \varphi_t^*(D_Y f).$$

于是

$$\begin{aligned}\frac{\varphi_t^*(D_Y f) - D_Y f}{t} &= \frac{D_{\varphi_t^*[Y]}(\varphi_t^*[f]) - D_Y(\varphi_t^*[f])}{t} - \frac{D_Y(\varphi_t^* f) - D_Y f}{t} \\ &= D_{\left\{\frac{\varphi_t^*[Y]-Y}{t}\right\}}\varphi_t^*[f] - D_Y\left(\frac{\varphi_t^* f - f}{t}\right).\end{aligned}$$

由于函数 $\varphi_t^*[f]$ 是一致可微的, 我们可以取 $t\to 0$ 时的极限, 得到

$$\begin{aligned}D_X(D_Y f) &= D_{D_X Y} f + D_Y(D_X f) \\ &= D_{[X,Y]} f + D_Y(D_X f).\end{aligned}$$

换句话说,

$$D_{[X,Y]} f = D_X(D_Y f) - D_Y(D_X f). \tag{6.13}$$

由于其定义是个导数, 则 $D_X Y$ 显然对 Y 是线性的:

$$D_X(aY_1 + bY_2) = aD_X Y_1 + bD_X Y_2,$$

其中 a 和 b 是常数, X 和 Y 是向量场. 由反对称性, 因而必对 X 也是线性的. 就是说,

$$D_{aX_1+bX_2}Y = [aX_1 + bX_2, Y] = a[X_1, Y] + b[X_2, Y] = aD_{X_1}Y + bD_{X_2}Y.$$

设 X, Y 为流形 M_2 上的向量场, ψ 为 M_1 到 M_2 上的微分同胚. 于是

$$\psi^*[X,Y] = [\psi^*X, \psi^*Y]. \tag{6.14}$$

事实上, 假设 X 生成流 φ. 于是

$$\begin{aligned}\psi^*[X,Y] = \psi^*D_XY = \psi^*\lim_{t\to 0}\left(\frac{\varphi_t^*Y - Y}{t}\right) &= \lim_{t\to 0}\frac{\psi^*\varphi_t^*Y - \psi^*Y}{t}\\ &= \lim_{t\to 0}\frac{\psi^*\varphi_t^*\psi^{-1*}\psi^*Y - \psi^*Y}{t}\\ &= \lim_{t\to 0}\frac{(\psi^{-1}\circ\varphi_t\circ\psi)^*\psi^*Y - \psi^*Y}{t}.\end{aligned}$$

因为 $\psi^{-1}\circ\varphi_t\circ\psi$ 是由 ψ^*X 生成的流, 因而得出这最后一个极限是 $D_{\psi^*X}\psi^*Y$, 从而证明了 (6.14) 式.

现在设 Y, Z 为 M 上的光滑向量场, X 为 φ 的无穷小生成元. 则

$$\begin{aligned}D_X[Y,Z] &= \lim_{t\to 0}\frac{\varphi_t^*[Y,Z] - [Y,Z]}{t}\\ &= \lim_{t\to 0}\frac{[\varphi_t^*Y, \varphi_t^*Z] - [Y,Z]}{t}\\ &= \lim_{t\to 0}\left\{\left[\frac{\varphi_t^*Y - Y}{t}, \varphi_t^*Z\right] + \left[Y, \frac{\varphi_t^*Z - Z}{t}\right]\right\}\\ &= [D_XY, Z] + [Y, D_XZ].\end{aligned}$$

因此

$$[X,[Y,Z]] = [[X,Y],Z] + [Y,[X,Z]] \tag{6.15}$$

由于李括号的反对称性, 等式 (6.15) 可以重写为

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0 \tag{6.16}$$

等式 (6.15) 或 (6.16) 以雅可比等式而知名.

9.7 线性微分形式

设 M为微分流形. 我们已对每个点 $x \in M$ 附加了一个向量空间 $T_x(M)$. 其对偶空间 $(T_x(M))^*$ 被称做 M 在 x 的余切空间, 并以 $T_x^*(M)$ 表示它. 因此 $T_x^*(M)$ 中的一个元素是 $T_x(M)$ 上的一个连续线性函数, 称之为余向量.

对 "连续的" 这个词给出某种诠释是适宜的. 在 M(因而 $T_x(M)$) 是有限维的情形, $T_x(M)$ 上的所有线性函数都是连续的, 不需进一步评说. 我们主要关心的也是这种情形. 更一般的情形, 设 l 为 $T_x(M)$ 上的线性函数. 对 x 的任何分图表 (U,α) 我们已将 $T_x(M)$ 等同于 V, 因而将 $\xi \in T_x(M)$ 等同于 $\xi_\alpha \in V$. 于是 l 决定了 V 上的一个线性函数 l_α, 其定义为

$$\langle \xi_\alpha, l_\alpha \rangle = \langle \xi, l \rangle \tag{7.1}$$

如果 (W,β) 是另一个分图表, 则

$$\langle \xi_\beta, l_\beta \rangle = \langle J_{\alpha\circ\beta^{-1}}(\beta(x))\xi_\beta, l_\alpha \rangle.$$

由于 $J_{\alpha\circ\beta^{-1}}(\beta(x))$ 是 V 到 V 的一个连续映射, 于是可以看出 l_α 为连续的充要条件是 l_β 为连续. 因此我们说, 如果 l_α 对某个 α(因而对任意的 α) 连续则 l 连续. 在这种情形中我们看出 (7.1) 式给出了将 $T_x^*(M)$ 与 V^* 的等同, 它把 l 变到 l_α. 上面最后一个等式说, 分图表的变化规则由

$$l_\beta = (J_{\alpha\circ\beta^{-1}}(\beta(x)))^* l_\alpha \tag{7.2}$$

给出.

设 f 是 M 上的可微函数, $x \in M$. 然后将每个 $\xi \in T_x(M)$ 送到 $\xi(f)$ 的 $T_x(M)$ 上的函数记作 $df(x)$. 于是

$$\langle \xi, df(x) \rangle = \xi f.$$

这容易看出 $df \in T_x^*(M)$. 事实上, 用一关于 x 的分图表 (U,α),

$$\langle \xi, df(x) \rangle = D_\xi(f_\alpha)(\alpha(x)).$$

注意, f 把每个元素 $df(x) \in T_x^*(M)$ 指派给了每个点 $x \in M$. 对每个 $x \in M$ 指定一个 $T_x^*(M)$ 中元素的映射被称为一个余向量场或者说一个线性微分形式. 由函数 f 决定的线性微分形式简单地记做 df.

设 ω 为一个线性微分形式. 于是对每个 $x \in M$ 有 $\omega(x) \in T_x^*(M)$. 设 $\mathfrak{a}$ 为 M 上一个总图表. 对每个 $(U,\alpha) \in \mathfrak{a}$ 我们得到一个 $\alpha(U)$ 上的 V^* - 值函数, 其定义为, 对 $v \in \alpha(U)$,

$$\omega_\alpha(v) = \{\omega(\alpha^{-1}(v))\}_\alpha, \tag{7.3}$$

如果 $(W,\beta) \in \mathfrak{a}$, 则 (7.2) 式表明, 对 $v \in \alpha(U \cap W)$ 有

$$\begin{aligned}\omega_\beta(\beta \circ \alpha^{-1}(v)) &= (J_{\alpha\circ\beta^{-1}}(\beta \circ \alpha^{-1}(v)))^* \omega_\alpha(v) \\ &= (J_{\beta\circ\alpha^{-1}}(v))^{-1*}\omega_\alpha(v).\end{aligned} \tag{7.4}$$

像前面一样, 等式 (7.4) 表明要求 ω 为光滑是有意义的. 我们说 ω 是个 C^k 微分形式是指如果对每个分图表 $(U,\alpha), \omega_\alpha$ 是个 V^* - 值的 C^k 函数. 由 (7.4) 式知, 只要对一个总图表的每个分图表进行验证即可. 又, 如果我们对每个 $(U,\alpha) \in \mathfrak{a}$, 在每个 $\alpha(U)$ 上定义了 V^* - 值函数 ω_α, 它们满足 (7.4) 式, 它们则通过 (7.3) 式定义了 M 上的一个线性微分形式 ω.

如果 ω 是个微分形式, f 为一函数, 我们定义 $f\omega$ 为 $f\omega(x) = f(x)\omega(x)$. 相似地, 我们定义 $\omega_1 + \omega_2$ 为

$$(\omega_1 + \omega_2)(x) = \omega_1(x) + \omega_2(x).$$

设 M_1 和 M_2 为微分流形, $\varphi : M_1 \to M_2$ 是个可微映射, 对每个 $x \in M_1$, 我们有映射 $\varphi_{*x} : T_x(M_1) \to T_{\varphi(x)}(M_2)$, 因此它定义了一个对偶映射

$$(\varphi_{*x})^* : T^*_{\varphi(x)}(M_2) \to T^*_x(M_1).$$

(读者可以验证, 如果 $l \in T^*_{\varphi(x)}(M_2)$, 则 $\xi \mapsto \langle\varphi_*(\xi), l\rangle$ 是 ξ 的连续线性函数, 这只要用一个分图表证实就可以了.)

现设 ω 为 M_2 上的微分形式. 它将 $\omega(\varphi(x)) \in T^*_{\varphi(x)}(M_2)$ 指派给了 $\varphi(x)$, 因而对每个 $x \in M_1$ 指定了一个元素 $(\varphi_{*x})^*(\omega(\varphi(x))) \in T^*_x(M_1)$. 因此我们将形式 ω “拉回”, 得到 M_1 上的一个形式, 我们称之为 $\varphi^*\omega$. 从而有

$$\varphi^*\omega(x) = (\varphi_{*x})^*\omega(\varphi(x)). \tag{7.5}$$

注意, φ^* 像函数时的情况那样, 不仅对微分同胚有定义而是对任何可微映射都有定义 (与向量场的情形形式对照).

容易得出用 M_1 上的 (U,α) 和 M_2 上的 (W,β) 的相容分图表表示的 φ^* 表达式. 事实上, 由 φ^* 的局部表达式我们看出, 对 $v \in \alpha(U)$ 有

$$(\varphi^*\omega)_\alpha(v) = (J_{\beta\circ\varphi\circ\alpha^{-1}}(v))^*\omega_\beta(\beta \circ \varphi \circ \alpha^{-1}(v)). \tag{7.6}$$

由 (7.6) 式我们知道 $\varphi^*\omega$ 在 ω 光滑时为光滑. φ^* 保持了代数运算也是清楚的:

$$\varphi^*(\omega_1+\omega_2)=\varphi^*\omega_1+\varphi^*\omega_2, \tag{7.7}$$

以及

$$\varphi^*(f\omega)=\varphi^*[f]\varphi^*(\omega). \tag{7.8}$$

如果 $\varphi:M_1\to M_2,\psi:M_2\to M_3$ 为可微映射, 则 (4.4) 和 (7.5) 式表明

$$(\psi\circ\varphi)^*\omega=\varphi^*\psi^*\omega. \tag{7.9}$$

设 $\psi:M_1\to M_2$ 为可微映射, f 为 M_2 上的可微函数. 于是 (4.6) 式和 df 的定义表明

$$d(\psi^*[f])=\psi^*df. \tag{7.10}$$

设 φ 为 M 上一个流, X 为其无穷小生成元, ω 为 M 上一个光滑线性微分形式. 于是形式 $\varphi_t^*\omega$ 局部有定义, 并且如同函数和向量场的情形那样, 当 $t\to 0$ 时

$$\frac{\varphi_t^*\omega-\omega}{t}$$

的极限存在. 我们可以用 (7.6) 式对此进行验算, 并就如同在向量场的情形那样进行. 此极限是个光滑的余向量场, 我们称之为 $D_X\omega$. 我们可用一个分图表来表达 $D_X\omega$, 完全像向量场那样去做.

如果 f 是可微函数, ω 为光滑微分形式, X 为 φ 的无穷小生成元, 则

$$D_X(f\omega)=D_X(f)\omega+fD_X\omega. \tag{7.11}$$

事实上,

$$\begin{aligned}D_X(f\omega)&=\lim_{t\to 0}\frac{\varphi_t^*f\omega-f\omega}{t}\\&=\lim_{t\to 0}\left(\frac{\varphi_t^*f-f}{t}\varphi_t^*(\omega)+f\frac{\varphi_t^*\omega-\omega}{t}\right)\\&=(D_Xf)\omega+fD_X\omega.\end{aligned}$$

如果 g 是 M 上的可微函数, 则

$$\frac{\varphi_t^*dg-dg}{t}=\frac{d\varphi_t^*g-dg}{t}=d\left(\frac{\varphi_t^*[g]-g}{t}\right).$$

借助于分图表的一个简单验算表明最后面的这个表达式的极限存在, 并且的确就是 $d(D_X g)$. 因此

$$D_X(df) = d(D_X f). \tag{7.12}$$

等式 (7.11) 和 (7.12) 表明, 如果

$$\omega = f_1 dg_1 + \cdots + f_k dg_k,$$

则

$$D_X\omega = (D_X f_1)dg_1 + \cdots + (D_X f_k)dg_k + f_1 d(D_X g_1) + \cdots + f_k d(D_X g_k). \tag{7.12$'$}$$

设 ω 是个光滑的线性微分形式, X 为光滑向量场. 于是 $\langle X,\omega\rangle$ 是个由

$$\langle X,\omega\rangle(x) = \langle X(x),\omega(x)\rangle$$

给出的光滑函数. 注意, $\langle X,\omega\rangle$ 对 X 和 ω 都是线性的. 也要注意到, 对任何光滑函数 f, 我们有

$$\langle X, df\rangle = D_X f. \tag{7.13}$$

9.8 用坐标计算

在本章的剩余部分中我们要假定所讨论的流形是有限维的. 设 M 是个微分流形, 其 $V = \mathbb{R}^n$. 如果 (U,α) 是 M 的一个分图表, 我们则可定义 U 上的函数 x_α^i 为

$$x_\alpha^i(x) = \alpha(x)\text{的第 } i \text{ 个坐标}. \tag{8.1}$$

如果 f 是 U 上的任一可微函数, 我们便可把等式 (9.1) 写成

$$f(x) = f_\alpha(x_\alpha^1(x), \cdots, x_\alpha^n(x)),$$

我们将其写做

$$f = f_\alpha(x_\alpha^1, \cdots, x_\alpha^n). \tag{8.2}$$

定义在 U 上的向量场 $\partial/\partial x_\alpha^i$ 为

$$\left(\frac{\partial}{\partial x_\alpha^i}\right)_\alpha (v) = \delta_i(= \prec 0, \cdots, \underset{\text{第 } i \text{ 个位置}}{1}, \cdots, 0 \succ). \tag{8.3}$$

如果 X 为 U 上任一向量场, 我们则有

$$X = X_\alpha^1 \frac{\partial}{\partial x_\alpha^1} + \cdots + X_\alpha^n \frac{\partial}{\partial x_\alpha^n}, \tag{8.4}$$

其中函数 X_α^i 的定义是

$$(X)_\alpha(\alpha(x)) = \prec X_\alpha^1(x), \cdots, X_\alpha^n(x) \succ . \tag{8.5}$$

等式 (8.4) 使我们将向量场 X 看作了一个"微分算子". 事实上, 由定义得到

$$D_X f = X_\alpha^1 \frac{\partial f_\alpha}{\partial x_\alpha^1} + \cdots + X_\alpha^n \frac{\partial f_\alpha}{\partial x_\alpha^n}. \tag{8.6}$$

因为 x_α^i 是 U 上的可微函数, dx_α^i 是 U 上的微分形式并且对所有 $v \in U$ 有

$$(dx_\alpha^i)_\alpha(v) = \delta_i. \tag{8.7}$$

特别地,

$$\left\langle \frac{\partial}{\partial x_\alpha^i}, dx_\alpha^j \right\rangle = \delta_i^j. \tag{8.8}$$

如果 ω 是 U 上一个微分形式, 则

$$\omega = a_{1\alpha} dx_\alpha^1 + \cdots + a_{n\alpha} dx_\alpha^n, \tag{8.9}$$

其中函数 $a_{i\alpha}$ 定义为

$$\omega_\alpha(\alpha(x)) = \prec a_{1\alpha}(x), \cdots, a_{n\alpha}(x) \succ \in \mathbb{R}^{n*}. \tag{8.10}$$

于是由定义推出

$$df = \frac{\partial f_\alpha}{\partial x_\alpha^1} dx_\alpha^1 + \cdots + \frac{\partial f_\alpha}{\partial x_\alpha^n} dx_\alpha^n. \tag{8.11}$$

等式 (8.11) 已将分图表变换下微分形式的转移法则结合为自己的一部分了. 事实上, 如果 (W, β) 是第二个分图表, 则在 $U \cap W$ 上我们由 (8.11) 式有

$$dx_\beta^i = \frac{\partial x_\beta^i}{\partial x_\alpha^1} dx_\alpha^1 + \cdots + \frac{\partial x_\beta^i}{\partial x_\alpha^n} dx_\alpha^n. \tag{8.12}$$

如果记 $\omega = a_{1\beta} dx_\beta^1 + \cdots + a_{n\beta} dx_\beta^n$, 使用替换 (8.12) 式我们得到

$$a_{j\alpha} = \sum \frac{\partial x_\beta^i}{\partial x_\alpha^j} a_{i\beta}.$$

现在

$$\left[\frac{\partial x_\beta^i}{\partial x_\alpha^j}\right]$$

是矩阵 $J_{\beta\circ\alpha^{-1}}$. 如果与 (8.10) 式进行比较, 我们看到我们已恢复了 (7.4) 式.

由于下标 α, β 等等把公式弄得杂乱不清, 我们便常常使用下面的记号约定: 用 x^i 的写法替代 x_α^i, 而用 y^j 去替代 x_β^j. 因此,

$$x^i = x_\alpha^i, y^j = x_\beta^j, z^k = x_\gamma^k, \text{等等}.$$

相似地, 我们将以 X^i 替代 X_α^i, Y^i 替代 X_β^i, a_i 替代 $a_{i\alpha}$, b_i 替代 $a_{i\beta}$, 等等.

于是等式 (8.1) 到 (8.12) 可以重写为

$$x^i(x) = \alpha(x)\text{的第 } i \text{ 个坐标}, \tag{8.1'}$$

$$f = f_\alpha(x^1, \cdots, x^n), \tag{8.2'}$$

$$\left(\frac{\partial}{\partial x^i}\right)_\alpha (v) = \delta_i, \tag{8.3'}$$

$$X = X^1 \frac{\partial}{\partial x^1} + \cdots + X^n \frac{\partial}{\partial x^n}, \tag{8.4'}$$

$$(X)_\alpha(\alpha(x)) = \prec X^1(x), \cdots, X^n(x) \succ \tag{8.5'}$$

$$D_X f = X^1 \frac{\partial f_\alpha}{\partial x^1} + \cdots + X^n \frac{\partial f_\alpha}{\partial x^n}, \tag{8.6'}$$

$$(dx^i)_\alpha(v) = \delta_i, \tag{8.7'}$$

$$\left\langle \frac{\partial}{\partial x^i}, dx^j \right\rangle = \delta_i^j, \tag{8.8'}$$

$$\omega = a_1 dx^1 + \cdots + a_n dx^n, \tag{8.9'}$$

$$\omega_\alpha(\alpha(x)) = \prec a_1(x), \cdots, a_n(x) \succ, \tag{8.10'}$$

$$df = \frac{\partial f_\alpha}{\partial x^1} dx^1 + \cdots + \frac{\partial f_\alpha}{\partial x^j} dx^j, \tag{8.11'}$$

$$dy^i = \frac{\partial y^i}{\partial x^1} dx^1 + \cdots + \frac{\partial y^i}{\partial x^n} dx^n, \tag{8.12'}$$

对 "拉回" 的公式也采用了简单的形式. 设 $\psi: M_1 \to M_2$ 为可微映射, 并设 M_1 为 m 维, M_2 为 n 维. 设 (U, α) 和 (W, β) 为相容分图表. 于是映射

$$\beta \circ \psi \circ \alpha^{-1} : \alpha(U) \to \beta(W)$$

由

$$y^i(\psi(x)) = y^i(x^1, \cdots, x^m), i = 1, \cdots, n \tag{8.13}$$

给出, 即 m 个实变量的 n 个函数. 如果 f 是 M_2 上的函数, 它在 W 上为

$$f = f_\beta(y^1, \cdots, y^n),$$

于是在 U 上

$$\psi^*[f] = f_\alpha(x^1, \cdots, x^m),$$

其中

$$f_\alpha(x^1, \cdots, x^m) = f_\beta(y^1(x^1, \cdots, x^m), \cdots, y^n(x^1, \cdots, x^m)). \tag{8.14}$$

“拉回”一个微分形式的规则也非常简易. 事实上, 如果

$$\omega = a_1 dy^1 + \cdots + a_n dy^n$$

为 ω 在 W 上的表达式, 则 $\psi^*\omega$ 在 U 上具有同一形式, 其中的这些 a 和 s, 我们现在把它们看成是 x 的函数, 并用 (8.12) 式展开它. 因此

$$\psi^*\omega = \sum a_i \frac{\partial y^i}{\partial x^j} dx^j,$$

其中 $a_i = a_i(y^1(x^1, \cdots, x^m), \cdots, y^n(x^1, \cdots, x^m))$.

设 $x \in U$. 于是

$$\psi_*\left(\frac{\partial}{\partial x^i}\right)(x) = \sum_j \frac{\partial y^j}{\partial x^i}(x) \frac{\partial}{\partial y^j}(\psi(x)) \tag{8.15}$$

给出了 ψ_{*x} 的公式.

习题

8.1 设 x, y 为 $\mathbb{E}^2$ 的直角坐标, (r, θ) 为 $\mathbb{E}^2 - \{0\}$ 上的极“坐标”. 用直角坐标表达向量场 $\frac{\partial}{\partial r}$ 和 $\frac{\partial}{\partial \theta}$. 又用极坐标表达 $\frac{\partial}{\partial x}$ 和 $\frac{\partial}{\partial y}$.

8.2 设 x, y, z 为 $\mathbb{E}^3$ 上的直角坐标. 设

$$X = y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}, Y = z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}, Z = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}.$$

计算 $[X, Y], [X, Z]$ 和 $[Y, Z]$. 注意, X 表示了绕 x 轴旋转的单参数群的无穷小生成元. 我们有时叫它为“绕 x 轴的无穷小旋转”. 对 Y 和 Z 也是相同的.

8.3 设

$$A = y\frac{\partial}{\partial x} + z\frac{\partial}{\partial y}, B = x\frac{\partial}{\partial z} + z\frac{\partial}{\partial x}, C = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}.$$

计算 $[A,B],[A,C]$ 和 $[B,C]$. 证明如果 $f(x,y,z) = x^2+y^2-z^2$, 则 $Af = Bf = Cf = 0$. 描述向量场 A, B 和 C 中每一个的积分曲线.

8.4 设

$$D = u\frac{\partial}{\partial v} + v\frac{\partial}{\partial u}, E = u\frac{\partial}{\partial v} - v\frac{\partial}{\partial u}, F = u\frac{\partial}{\partial u} - v\frac{\partial}{\partial v}.$$

计算 $[D,E],[D,F]$ 和 $[E,F]$.

8.5 设 $P_1,\cdots,P_n$ 为 $x^1,\cdots,x^n$ 的多项式, 它们均无常数项, 即

$$P_i(0,\cdots,0) = 0.$$

设

$$I = x^1\frac{\partial}{\partial x^1} + \cdots + x^n\frac{\partial}{\partial x^n}$$

和

$$X = P_1\frac{\partial}{\partial x^1} + \cdots + P_n\frac{\partial}{\partial x^n}$$

证明

$$[I,X] = 0$$

当且仅当 P_i 为线性的.[提示: 考虑将这些 P_i 展开为齐次项.]

8.6 设 X 和 P_i 同于习题 8.5 的, 并假定 P_i 都是线性的, 又设

$$A = \lambda_1 x^1\frac{\partial}{\partial x^1} + \cdots + \lambda_n x^n\frac{\partial}{\partial x^n},$$

并假定当 $i \neq j$ 时

$$\lambda_i \neq \lambda_j.$$

证明 $[A,X] = 0$ 当且仅当 $P_i = \mu_i x^i$, 即对某些 $\mu^1,\cdots,\mu^n$

$$X = \mu_1 x^1\frac{\partial}{\partial x^1} + \cdots + \mu_n x^n\frac{\partial}{\partial x^n}.$$

8.7 设 A 同于习题 8.6, 并假设另外还有对任意 i,j,r 有

$$\lambda_i \neq \lambda_j + \lambda_r$$

证明如果 P_i 最多为二次, 则

$$[A,X] = 0$$

的充要条件是 $P_i = \mu_i x^i$. 将此结果推广到 P_i 最多为 m 次多项式的情形.

9.9 黎曼度量

设 M 为有限维的微分流形. M 上的一个黎曼度量m 是一个规则, 它对每个 $x \in M$ 的向量空间 $T_x(M)$ 指定一个正定的内积 $(\ ,\ )_{m,x}$. 只要从上下文中已经明白, 通常我们去掉下标 m 和 x. 因此, 如果 m 是 M 上的一个黎曼度量, $x \in M, \xi, \eta \in T_x(M)$, 我们记 ξ 和 η 的内积为

$$(\xi, \eta) = (\xi, \eta)_{m,x}.$$

设 (U, α) 为 M 的一个分图表. 定义 U 上函数 g_{ij} 为

$$g_{ij}(x) = \left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x)\right), \tag{9.1}$$

故 $g_{ij} = g_{ji}$. 如果 $\xi, \eta \in T_x(M)$ 为

$$\xi = \sum \xi^i \frac{\partial}{\partial x^i}(x), \eta = \sum \eta^i \frac{\partial}{\partial x^i}(x),$$

则

$$(\xi, \eta) = \sum_{i,j} g_{ij}(x) \xi^i \eta^j.$$

因为 $dx^1(x), \cdots, dx^n(x)$ 是 $T_x^*(M)$ 的一组对偶于

$$\frac{\partial}{\partial x^1}(x), \cdots, \frac{\partial}{\partial x^n}(x)$$

的基, 所以我们有

$$\xi^i = \langle \xi, dx^i(x) \rangle, \quad \eta^j = \langle \eta, dx^j(x) \rangle,$$

故上面最后的等式可以写成

$$(\xi, \eta)_{m,x} = \sum g_{ij}(x) \langle \xi, dx^i \rangle \langle \eta, dx^j \rangle. \tag{9.2}$$

等式 (9.2) 通常被写成更简明的形式:

$$m \restriction U = \sum g_{ij}(x) dx^i dx^j. \tag{9.3}$$

[(9.3) 式可以被解释为 (9.2) 式的一种缩写方式.]

设 (W, β) 为第二个分图表, 并对 $x \in W$ 有

$$h_{kl}(x) = \left(\frac{\partial}{\partial y^k}(x), \frac{\partial}{\partial y^l}(x)\right),$$

即

$$m \restriction W = \sum h_{kl} dy^k dy^l. \tag{9.4}$$

于是对 $x \in U \cap W$ 我们有

$$\frac{\partial}{\partial x^i}(x) = \sum \frac{\partial y^k}{\partial x^i}(x) \frac{\partial}{\partial y^k}(x), \quad \frac{\partial}{\partial x^j}(x) = \sum \frac{\partial y^l}{\partial x^j}(x) \frac{\partial}{\partial y^l}(x),$$

故

$$g_{ij}(x) = \left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x) \right) = \sum_{k,l} \frac{\partial y^k}{\partial x^i}(x) \frac{\partial y^l}{\partial x^j}(x) h_{kl}(x),$$

即

$$g_{ij} = \sum_{k,l} h_{kl} \frac{\partial y^k}{\partial x^i} \frac{\partial y^l}{\partial x^j}. \tag{9.5}$$

注意, 如果我们形式地将 (8.12) 式代入 (9.4) 式中的各个 dy, 并合并 $dx^i dx^j$ 的系数, 我们得到的结果便会是 (9.5) 式.

无论怎样, 由 (9.5) 式可清楚知道, 如果 h_{ij} 全都是 W 上的光滑函数, 则 g_{ij} 在 $U \cap W$ 上光滑. 由于这个原因, 我们在 (9.3) 式给出的函数 g_{ij} 对任何属于 M 上一个总图表 $\mathfrak{a}$ 的分图表 (U, α) 为光滑时便说这个黎曼度量是光滑的. 另外, 如果我们有对每个 $(U, \alpha) \in \mathfrak{a}$ 上定义的函数 $g_{ij} = g_{ji}$, 使得

(i) $\sum g_{ij} \xi^i \xi^j > 0$, 除了对所有 $x \in U$ 有 $\xi = 0$ 外.

(ii) 转移法则 (9.5) 式成立,

则 g_{ij} 定义了 M 上的一个黎曼度量. 在后面的讨论中我们假定所有的黎曼度量是光滑的.

设 $\psi : M_1 \to M_2$ 为可微映射, m 为 M_2 上的一个黎曼度量. 对任何一个 $x \in M_1$, 定义在 $T_x(M_1)$ 上的 $(,)_{\psi^*(m),x}$ 为

$$(\xi, \eta)_{\psi^*(m),x} = (\psi_*(\xi), \psi_*(\eta))_{m,\psi(x)}. \tag{9.6}$$

它定义了一个 ξ 和 η 的对称双线性函数. 然而, 由于可以有 $\xi \neq 0$ 而 $\psi_*(\xi) = 0$ 的情形, 故它不必是正定的. 因此, 一般说来, (9.6) 式没有定义出 M_1 上的一个黎曼度量. 但对某些 ψ, 这确实是个黎曼度量.

称一个可微映射 $\psi : M_1 \to M_2$ 是个浸入是说如果对所有 $x \in M_1, \psi_{*,x}$ 是个单射 (即一一的).

如果 $\psi : M_1 \to M_2$ 为浸入, m 为 M_2 上的黎曼度量, 则由 (9.6) 式, 我们在 M_1 上定义了一个黎曼度量 $\psi^*(m)$.

设 (U,α) 和 (W,β) 是 M_1 和 M_2 上的相容分图表, 并设

$$m \restriction W = \sum h_{kl} dy^k dy^l.$$

于是

$$\psi^*(m) \restriction U = \sum g_{ij} dx^i dx^j.$$

其中

$$\begin{aligned} g_{ij}(x) &= \left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x)\right)_{\psi^*(m),x} \\ &= \left(\psi_*\left(\frac{\partial}{\partial x^i}(x)\right), \psi_*\left(\frac{\partial}{\partial x^j}(x)\right)\right)_{m,\psi(x)} \\ &= \left(\sum \frac{\partial y^k}{\partial x^i}\frac{\partial}{\partial y^k}(\psi(x)), \sum \frac{\partial y^l}{\partial x^j}\frac{\partial}{\partial y^l}(\psi(x))\right)_{m,\psi(x)} \\ &= \sum_{k,l} h_{kl}\frac{\partial y^k}{\partial x^i}\frac{\partial y^l}{\partial x^j} \end{aligned}$$

这又正好是 (9.5) 式 (但有不同的解释). 或者更简约地,

$$\psi^*(m) \restriction U = \sum \psi^*(h_{kl})\psi^*(dy^k)\psi^*(dy^l).$$

我们来给出这些公式的一些例子. 如果 $M = \mathbb{R}^n$, 则等同图表在 $\mathbb{R}^n$ 上诱导了一个由

$$(dx^1)^2 + \cdots + (dx^n)^2$$

给出的黎曼度量. 让我们看一看在极坐标下 $\mathbb{R}^2$ 和 $\mathbb{R}^3$ 上它是什么样的.

在 $\mathbb{R}^2$ 中如果我们记

$$x^1 = r\cos\theta,\ x^2 = r\sin\theta,$$

则

$$\begin{aligned} dx^1 &= \cos\theta dr - r\sin\theta d\theta \\ dx^2 &= \sin\theta dr + r\cos\theta d\theta, \end{aligned}$$

故

$$(dx^1)^2 + (dx^2)^2 = dr^2 + r^2 d\theta^2 \tag{9.7}$$

注意, 只要形式 dr 和 $d\theta$ 有定义, 即在整个 $\mathbb{R}^2-\{0\}$ 上有意义, 则 (9.7) 式就成立.(尽管函数 θ 不是在整个 $\mathbb{R}^2-\{0\}$ 上有定义时, 形式 $d\theta$ 仍可在此上有定义. 事实上, 我们可以写成

$$d\theta=\frac{x^1dx^2-x^2dx^1}{(x^1)^2+(x^2)^2}.) \tag{9.8}$$

在 $\mathbb{R}^3$ 中, 我们引进

$$\begin{aligned} x^1&=r\cos\varphi\sin\theta,\\ x^2&=r\sin\varphi\sin\theta,\\ x^3&=r\cos\theta. \end{aligned}$$

于是

$$\begin{aligned} dx^1&=\cos\varphi\sin\theta dr-r\sin\varphi\sin\theta d\varphi+r\cos\varphi\cos\theta d\theta,\\ dx^2&=\sin\varphi\sin\theta dr+r\cos\varphi\sin\theta d\varphi+r\sin\varphi\cos\theta d\theta,\\ dx^3&=\cos\theta dr-r\sin\theta d\theta. \end{aligned}$$

因此

$$(dx^1)^2+(dx^2)^2+(dx^3)^2=dr^2+r^2\sin^2\theta d\varphi^2+r^2d\theta. \tag{9.9}$$

另外, 一旦右端的形式被定义则 (9.9) 式就成立, 而此时表明的是 $(x^1)^2+(x^2)^2\neq 0$ 的情形.

让我们考虑单位球的映射 $\iota: S^2\to\mathbb{R}^3$, 它由将 S^2 的点看作 $\mathbb{R}^3$ 中点组成. 于是我们得到 S^2 上一个诱导的黎曼度量.

令

$$\overline{d\theta}=\iota^*d\theta, \overline{d\varphi}=\iota^*d\varphi,$$

故形式 $\overline{d\theta}$ 和 $\overline{d\varphi}$ 在 $U=S^2-\{\prec 0,0,1\succ,\prec 0,0,-1\succ\}$ 有定义. 于是在 U 上我们可以写成 (由于在 S^2 上 $r=1$)

$$\iota^*((dx^1)^2+(dx^2)^2+(dx^3)^2)=\overline{d\theta}^2+\sin^2\theta\overline{d\varphi}^2. \tag{9.10}$$

我们回到对一般情形的思考. 设 M 为微分流形, $C:I\to M$ 为微分映射, 其中 I 为 $\mathbb{R}^1$ 中的区间. 设 t 表示 I 上等同分图表的坐标. 令

$$C'(s)=C_*\left(\frac{\partial}{\partial t}\right)(s), s\in I,$$

故 $C'(s) \in T_{C(s)}M$ 为曲线 C 在 s 的切向量. 如果 (U,α) 是 M 的一个分图表, $x^1,\cdots,x^n$ 为 (U,α) 的坐标函数, 于是, 如果对某个 $I' \subset I$ 有 $C(I') \subset U$,

$$\alpha \circ C = \prec x^1 \circ C, \cdots, x^n \circ C \succ,$$

那么

$$C'(t)_\alpha = \prec \frac{dx^1 \circ C}{dt}, \cdots, \frac{dx^n \circ C}{dt} \succ (t).$$

在不会发生混淆时, 我们将略去 $\circ C$ 而简单地写为

$$\alpha \circ C(t) = \prec x^1(t), \cdots, x^n(t) \succ \text{和} C'(t)_\alpha = \prec x^{1'}(t), \cdots, x^{n'}(t) \succ .$$

现在设 m 为 M 上的一个黎曼度量. 则 $\|C'(t)\| = (C'(t), C'(t))^{1/2}$ 是一个连续函数. 事实上, 借助于一个分图表, 我们可以写为

$$\|C'(t)\| = \sqrt{\sum g_{ij}(C(t))x^{i'}(t)x^{j'}(t)}.$$

积分

$$\int_I \|C'(t)\| dt \tag{9.11}$$

被称做曲线 C 的长. 如果 $\|C'(t)\|$ 在 I 上可积, 它就有定义. 例如, 如果 I 和 $\|C'(t)\|$ 都为有界就确属此情形. 注意, 长度与参数化无关. 更准确地说, 设 $\varphi: J \to I$ 为一一的可微映射, 设 $C_1 = C \circ \varphi$, 则对任何 $\tau \in J$ 我们有

$$C_1' = C_{1*}\left(\frac{\partial}{\partial \tau}\right) = C_* \circ \varphi_*\left(\frac{\partial}{\partial \tau}\right) = \frac{dt}{d\tau} C_*\left(\frac{\partial}{\partial t}\right),$$

即

$$C_1'(\tau) = \frac{dt}{d\tau} C'(t).$$

因此

$$\|C_1'(\tau)\| = \|C'(\varphi(\tau))\| \left|\frac{d\varphi}{d\tau}\right|. \tag{9.12}$$

另一方面, 由积分的换元公式知道, (9.12) 式表明

$$\begin{aligned}\int_I \|C'(\cdot)\| &= \int_J \|C'(\varphi(\cdot))\| \left|\frac{d\varphi}{d\tau}\right| \\ &= \int_J \|C_1'(\cdot)\|.\end{aligned}$$

更为一般地，我们称定义在区间 I 上的一条曲线 C 是逐段可微的是说如果

(i) C 连续；

(ii) $I = I_1 \cup \cdots \cup I_r$，而 C 在每个 I_j 上是另一个定义.

在严格包含了 I_j 的某个区间 I_j' 上的可微曲线的限制.(因此一条逐段可微曲线是一条具有有限个"拐角"的曲线.) 如果 C 逐段可微，则 $\|C'(t)\|$ 除去有限个 t 值外是连续的，而在那些有限个点上它可能会有一个跳跃的不连续性. 特别，积分 (9.11) 式存在而所有这种曲线均有长度.

习题

设 C 为一曲线，它将 I 一一地映成 $\mathbb{R}^n$ 中一条直线段. 证明 C 的长与此线段的长相同.

设 $C:[0,1]\to\mathbb{R}^2$ 为 $C(0)=0, C(1)=v\in\mathbb{R}^2$ 的曲线. 如果我们使用表达式 (9.7)，则有

$$\begin{aligned}\int \|C'(t)\|dt &= \int \sqrt{(r'(t))^2+(r(t)\theta'(t))^2}\,dt\\ &\geqslant \int |r'(t)|dt\\ &\geqslant \int_0^1 r'(t)dt\\ &= \|v\|,\end{aligned}$$

等号成立当且仅当 $\theta'\equiv 0$ 且 $r'\geqslant 0$. 因此在所有联结 0 到 v 的曲线中，直线段具有最短长度.

相似地，在球面上，设 C 为任意曲线 $C:[0,1]\to S^2$，满足 $C(0)=(0,0,1)$ 和 $C(1)=p\neq(0,0,-1)$，并设 $\theta_1=\theta(C(1))$. 于是

$$\begin{aligned}\int \|C'(t)\|dt &= \int \sqrt{(\theta'(t))^2+\sin^2\theta(\varphi'(t))^2}\,dt\\ &\geqslant \int_0^1 |\theta'(t)|dt.\end{aligned}$$

如果设 t_1 表示 $[0,1]$ 中第一个使 $\theta=\theta_1$ 的点，则

$$\int \|C'(t)\|dt \geqslant \int_0^1 |\theta'(t)|\,dt = \int_0^{t_1} |\theta'(t)|\,dt \geqslant \int_0^{t_1}\theta'(t) = \theta_1,$$

使等号成立只有 $\varphi'\equiv 0$ 和 $t_1=1$. 因此联结 S^2 中任两点的最短曲线是联结它们的大圆.

在上面两个例子中我们都得助于一个非常幸运的坐标选择 (平面中的极坐标和球面上的类似极坐标). 我们将在第十三章的 11 节中看到这不是偶然的. 我们会看到在任何黎曼流形上人们可引进容易描述具局部极小化长度曲线的局部坐标.

第十章　流形上的积分学

在这一章里我们要研讨流形上的积分. 为了阐述积分, 我们不得不限制所考虑的流形的类别. 本章中我们将假定所有出现的流形 M 均满足下面两个条件:

(1) M 是有限维的.

(2) M 具有一个总图表 $\mathfrak{a}$, 它包含了 (最多) 可数个分图表, 即 $\mathfrak{a} = \{(U_i, \alpha_i)\}_{i=1,2,\cdots}$

在我们真正开始做积分之前, 要建立好几个技术层面上的东西. 前两节所做的就是这些.

10.1　紧性

流形 M 的一个子集 A 被称作紧的是说它有如下性质:

(i) 如果有任一组开集 $\{U_\iota\}$, 满足

$$A \subset \bigcup_\iota U_\iota,$$

则存在有限个 U_ι , 譬如 $U_{\iota_1}, \cdots, U_{\iota_r}$, 使得

$$A \subset U_{\iota_1} \cup \cdots \cup U_{\iota_r},$$

换一种说法就是:

(ii) 集合 A 为紧的充要条件是对任意闭集的族 $\{F_\iota\}$, 使得

$$A \cap \bigcap_\iota F_\iota = \phi,$$

则存在有限个 F_ι 使得

$$A \cap F_{\iota_1} \cap \cdots \cap F_{\iota_r} = \phi.$$

(i) 和 (ii) 的等价性由取 U_ι 为 F_ι 的补集便可得到.

第四章 4.5 节中我们曾建立过这样的事实: 如果 $M = U$ 是 $\mathbb{R}^n$ 中开集, 则 $A \subset U$ 为紧集的充要条件是 A 为 $\mathbb{R}^n$ 中的一个闭的有界子集.

关于紧性我们做出一些进一步的然而是显而易见的评注:

(iii) 如果 $A_1, \cdots, A_r$ 为紧, 则 $A_1 \cup \cdots \cup A_r$ 也为紧.

事实上, 如果 $\{U_\iota\}$ 覆盖 $A_1 \cup \cdots \cup A_r$, 它必覆盖每个 A_j. 因而我们可对每个 j 选取覆盖 A_j 的有限个元的子组. 这些子组的并构成了覆盖 $A_1 \cup \cdots \cup A_r$ 的一个有限子组.

(iv) 如果 $\psi : M_1 \to M_2$ 连续, $A \subset M_1$ 为紧集, 则 $\psi[A]$ 为紧.

事实上, 如果 $\{U_\iota\}$ 覆盖了 $\psi[A]$, 则 $\{\psi^{-1}(U_\iota)\}$ 覆盖 A. 如果 U_ι 为开则因 ψ 连续, 有 $\psi^{-1}(U_\iota)$ 为开. 故可选取 $\iota_1, \cdots, \iota_r$ 使得

$$A \subset \psi^{-1}(U_{\iota_1}) \cup \cdots \cup \psi^{-1}(U_{\iota_r}),$$

它意味着 $\psi[A] \subset U_{\iota_1} \cup \cdots \cup U_{\iota_r}$.

从这个性质我们看出, 如果 $A = A_1 \cup \cdots \cup A_r$, 其中每个 A_j 包含在某个 W_i 中, 而 (W_i, β_i) 是个分图表, 并且 $\beta_i(A_j)$ 是 $\mathbb{R}^n$ 中的一个紧子集, 则 A 是个紧集. 特别, 流形 M 自身可以是紧的. 例如, 我们可将 S^n 写成上下两个半球面的并: $S^n = \{x : x^{n+1} \geqslant 0\} \cup \{x : x^{n+1} \leqslant 0\}$. 每个半球面是紧的. 事实上, 上半球由 8.1 节的映射 φ_1 映成了 $\{y : \|y\| \leqslant 1\}$, 而下半球由 φ_2 映成同样的集合. 因此球面是紧的.

另一方面, $\mathbb{R}^n$ 中的开子集是不紧的. 但是, 它可以表示为可数个紧集的并. 事实上, 如果 $U \subset \mathbb{R}^n$ 为开集, 令

$$A_n = \{x \in U : \|x\| \leqslant n \quad 而 \quad \rho(x, \partial U) \geqslant \frac{1}{n}\}.$$

容易验证 A_n 是紧的, 且 $\cup A_n = U$.

由于有条件 (2), 我们对任意所讨论的流形 M 可以说同样的话:

命题 1.1 任何满足 (1) 和 (2) 的流形 M 可以写成

$$M = \bigcup_{i=1}^{\infty} A_i,$$

其中每个 $A_i \subset M$ 为紧集.

证明 事实上, 由 (2) 有

$$M = \bigcup_{j=1}^{\infty} U_j,$$

并且由上面的讨论知, 每个 U_j 可以表示为紧集的一个可数并. 由于可数并的可数并仍是可数的, 便有此命题. □

直接的推论是:

命题 1.2 设 M 为 [满足 (1) 和 (2) 的] 流形, $\{U_\iota\}$ 为 M 的开覆盖. 则我们可选取一个可数的子组 $\{U_j\}$ 使得

$$\cup U_j = M.$$

证明 记 $M = \cup A_r$, 其中每个 A_r 为紧. 对每个 r, 可选取有限个 $U_{r,1}, U_{r,2}, \cdots, U_{r,k_r}$, 使得

$$A_r \subset U_{r,1} \cup \cdots \cup U_{r,k_r}.$$

汇集

$$\{U_{r,k}\}_{\substack{r=1,\cdots,\infty \\ k=1,\cdots,k_r}}$$

是个覆盖了 M 的可数子组. □

10.2 1 的分解

在后面的讨论中, 如果有一个把函数, 向量场等等能 "分裂" 成 "小片段" 的方法, 则对我们来说会很方便. 为此目的, 我们引进下面的记号:

定义 2.1 一组 C^∞ 函数 $\{g_i\}$ 被称作为一个 1 的分解是说, 如果成立:

(i) 对所有 $i, g_i \geqslant 0$;

(ii) 对所有 i, supp g_i 为紧集; [1]

(iii) 每个 $x \in M$ 有一个邻域 V_x, 使得除去有限个外对所有其他的 i 成立 $V_x \cap \operatorname{supp} g_i = \varnothing$;

(iv) 对所有 $x \in M, \sum g_i(x) = 1$.

注意, 由于有 (iii), 在 (iv) 中出现的那个和式实际上是有限的, 其原因在于对任一个 x, 除去有限个 $g_i(x)$ 外全都为零. 另外还要注意有:

命题 2.1 如果 A 是个紧集, $\{g_i\}$ 是 1 的一个分解, 则对除去有限个 i 外, 对所有 i 有

$$A \cap \operatorname{supp} g_i = \varnothing.$$

证明 事实上, 每个 $x \in A$ 都有一个由 (iii) 给出的邻域 V_x. 集合 $\{V_x\}_{x \in A}$ 是 A 的一个开覆盖. 因为 A 为紧, 故可选取有限个子组 $\{V_1, \cdots, V_r\}$, 使 $A \subset V_1 \cup \cdots \cup V_r$. 因为每个 V_k 只与有限个 supp g_i 有非空交集, 故它们的并, 从而 A 也只与有限个 supp g_i 有非空交集. □

定义 2.2 设 $\{U_\iota\}$ 为 M 的一组开覆盖, $\{g_j\}$ 为 1 的分解. 称 $\{g_j\}$ 从属于 $\{U_\iota\}$ 是说如果对每个 j, 存在一个 $\iota(j)$ 使得

$$\operatorname{supp} g_j \subset U_{\iota(j)}. \tag{2.1}$$

定理 2.1 设 $\{U_\iota\}$ 是 M 的任一组开覆盖. 于是存在一个从属于 $\{U_\iota\}$ 的一个 1 的分解 $\{g_j\}$.

我们在下面所给出的证明是 Bonic 和 Frampton 给出的. [2]

首先介绍一些预备性的概念.

由

$$f(u) = \begin{cases} e^{-1/u}, & \text{如果 } u > 0, \\ 0, & \text{如果 } u \leqslant 0. \end{cases}$$

定义的 $\mathbb{R}$ 上的函数 f 是 C^∞ 的. 对于 $u \neq 0$, 显然 f 具有所有阶的导数. 要验证 f 在 0 是 C^∞, 只要证明当 u 从右边趋向 0 时 $f^{(k)}(u) \to 0$ 即可. 然而 $f^{(k)}(u) = P_k(1/u) \cdot e^{-1/u}$, 其中 P_k 是个 $2k$ 次的多项式. 因 e^s 趋向无穷大快于

[1] 回忆, supp g 是集合 $\{x : g(x) \neq 0\}$ 的闭包.

[2] Smooth function on Banach manifolds, J.Math and Mech. **15**, 877~898(1966)

任何多项式,

$$\lim_{u\to 0}f^{(k)}(u) = \lim_{s\to\infty}P_k(s)e^{-s} = 0.$$

注意, $f(u) > 0$ 当且仅当 $u > 0$. 现考虑在 $\mathbb{R}$ 上的函数 g_a^b, 其定义为

$$g_a^b(x) = f(x-a)f(b-x).$$

于是, g_a^b 为 C^∞, 非负, 且

$$g_a^b(x) \neq 0 \quad \text{当且仅当} \quad a < x < b.$$

更一般地, 如果 $\boldsymbol{a} = \prec a^1, \cdots, a^k \succ$ 和 $\boldsymbol{b} = \prec b^1, \cdots, b^k \succ$, 定义在 $\mathbb{R}^k$ 上函数 $g_{\boldsymbol{a}}^{\boldsymbol{b}}$ 为

$$g_{\boldsymbol{a}}^{\boldsymbol{b}}(x) = g_{a^1}^{b^1}(x^1)g_{a^2}^{b^2}(x^2)\cdots g_{a^k}^{b^k}(x^k),$$

其中 $x = \prec x^1, \cdots, x^k \succ$. 于是, $g_{\boldsymbol{a}}^{\boldsymbol{b}} \geqslant 0, g_{\boldsymbol{a}}^{\boldsymbol{b}} \in C^\infty$ 且

$$g_{\boldsymbol{a}}^{\boldsymbol{b}}(x) > 0 \quad \text{当且仅当} \quad a^1 < x^1 < b^1, \cdots, a^k < x^k < b^k. \tag{2.2}$$

引理 2.1 设 $f_1, \cdots, f_k$ 为流形 M 上的 C^∞ 函数, $W = \{x : a^1 < f_1(x) < b^1, \cdots, a^k < f_k(x) < b^k\}$. 于是存在一个非负 C^∞ 函数 g, 使得 $W = \{x : g(x) > 0\}$.

事实上, 如果定义 g 为

$$g(x) = g_{\boldsymbol{a}}^{\boldsymbol{b}}(f_1(x), \cdots, f_k(x)),$$

则 g 显然具有所要求的性质.

我们转向定理 2.1 的证明.

证明 对每个 $x \in M$, 选取 U_ι 包含 x, 又设 x 的一个分图表为 (U, α). 那么, $\alpha(U \cap U_\iota)$ 是 $\mathbb{R}^n$ 中包含 $\alpha(x)$ 的一个开集. 选择 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 使

$$\alpha(x) \in \mathrm{int}\square_{\boldsymbol{a}}^{\boldsymbol{b}}, \quad \overline{\square_{\boldsymbol{a}}^{\boldsymbol{b}}} \subset \alpha(U \cap U_\iota).$$

设 $W_x = \alpha^{-1}(\mathrm{int}\square_{\boldsymbol{a}}^{\boldsymbol{b}})$. 于是

$$\overline{W_x} \subset U_\iota \quad \text{且} \quad \overline{W_x}\text{为紧集}. \tag{2.3}$$

又如果 $x^1, \cdots, x^n$ 为 α 给出的坐标,

$$W_x = \{y : a^1 < x^1(y) < b^1, \cdots, a^n < x^n(y) < b^n\}.$$

由上面引理, 我们可以找到一个非负的 C^∞ 函数 f_x 使得

$$W_x = \{y : f_x(y) > 0\}.$$

因为 $x \in W_x$ 所以 $\{W_x\}$ 覆盖了 M. 命题 1.2 使我们能选取一个可数子覆盖 $\{W_i\}$. 我们用 f_i 代表相对应的函数, 就是说, 如果 $W_i = W_x$, 则令 $f_i = f_x$.

设

$$\begin{aligned}
V_1 &= W_1 = \{x : f_1(x) > 0\},\\
V_2 &= \{x : f_2(x) > 0, f_1(x) < \frac{1}{2}\},\\
&\vdots\\
V_r &= \{x : f_r(x) > 0, f_1(x) < 1/r, \cdots, f_{r-1}(x) < 1/r\}.
\end{aligned}$$

显然 V_j 为开集且 $V_j \subset W_j$, 那么, 由 (2.3) 式有

$$\overline{V_j}\text{为紧集}\quad\text{且}\quad \overline{V_j} \subset U_\iota, \tag{2.4}$$

其中 $\iota = \iota(j)$ 是某个指标.

对每个 $x \in M$, 令 $q(x)$ 代表第一个使 $f_q(x) > 0$ 的整数 q. 因而如果 $p < q(x)$ 则 $f_p(x) = 0$, 而 $f_{q(x)}(x) > 0$.

设 $V_x = \{y : f_{q(x)}(y) > \frac{1}{2} f_{q(x)}(x)\}$. 因为 $f_{q(x)}(x) > 0$, 从而得到 $x \in V_x$, 且 V_x 为开集. 进一步有

$$V_x \cap \overline{V_r} = \varnothing, \text{其中} r > q(x), \frac{1}{r} < \frac{1}{2} f_{q(x)}(x). \tag{2.5}$$

按照上面引理, 每个 V_i 由 $V_i = \{x : \overline{g}_i(x) > 0\}$ 给出, 其中 $\overline{g}_i$ 为某个 C^∞ 函数. 设 $g = \sum \overline{g}_i$. 由 (2.5) 式, 在每个 x 的邻域内这确实是个有限和. 因此 g 为 C^∞. 因为 $x \in V_{q(x)}, \overline{g}_{q(x)}(x) > 0$. 于是 $g > 0$. 令

$$g_j = \overline{g}_j / g,$$

我们断定 $\{g_j\}$ 就是所想要的 1 的分解. 事实上, 由构造知 (i) 成立, (ii) 及 (2.1) 式由 (2.4) 式推出, (iii) 由 (2.5) 式得到, 而 (iv) 因构造而成立. □

10.3 密度

如果我们把 $\mathbb{R}^n$ 看作是一个微分流形, 则对积分的换元法则表明, 被积项没有像在分图表变化下的一个函数的同样的转移法则. 由于这个原因, 我们不能期望在流形上对函数进行积分. 我们现在要引进我们可以进行积分的那类对象.

定义 3.1 *一个密度 ρ 是一个规则, 它对每个 M 上分图表 (U,α) 指定一个定义在 $\alpha(U)$ 上的函数 ρ_α, 它满足于下面的转移法则: 如果 (W,β) 是 M 上第二个分图表, 则对 $v \in \alpha(U \cap W)$ 有*

$$\rho_\alpha(v) = \rho_\beta(\beta \circ \alpha^{-1}(v))|\det J_{\beta\circ\alpha^{-1}}(v)|. \tag{3.1}$$

如果 $\mathfrak{a}$ 是 M 上一个总图表, 并对每个 $(U_i,\alpha_i) \in \mathfrak{a}$ 给出函数 ρ_{α_i}, 它们满足 (3.1) 式, 于是 ρ_{α_i} 定义了 M 上一个密度 ρ. 事实上, 如果 (U,α) 是 M 上一个任意的分图表 (不必属于 $\mathfrak{a}$), 我们定义 ρ_α 为, 对 $v \in \alpha(U \cap U_i)$ 有

$$\rho_\alpha(v) = \rho_{\alpha_i}(\alpha_i \circ \alpha^{-1}(v))|\det J_{\alpha_i\circ\alpha^{-1}}(v)|.$$

这个定义是相容的: 如果 $v \in \alpha(U \cap U_i) \cap \alpha(U \cap U_j)$, 则由 (3.1) 式, 有

$$\begin{aligned}
&\rho_{\alpha_j}(\alpha_j \circ \alpha^{-1}(v))\left|\det J_{\alpha_j\circ\alpha^{-1}}(v)\right| \\
&\quad = \rho_{\alpha_i}(\alpha_i \circ \alpha_j^{-1}(\alpha_j \circ \alpha^{-1}(v)))\left|\det J_{\alpha_i\circ\alpha_j^{-1}}(\alpha_j \circ \alpha^{-1}(v))\right|\left|\det J_{\alpha_j\circ\alpha^{-1}}(v)\right| \\
&\quad = \rho_{\alpha_i}(\alpha_i \circ \alpha^{-1}(v))\left|\det J_{\alpha_i\circ\alpha^{-1}}(v)\right|,
\end{aligned}$$

其中运用了链规则和行列式的乘法性质.

鉴于 (3.1) 式, 谈及密度的局部光滑性质是有意义的. 我们称密度 ρ 是 C^k 的是说, 如果对任意分图表 (U,α) 函数 ρ_α 是 C^k 的. 照例, 只要证明所有分图表 (U,α) 属于某一总图表即可. 相似地, 我们说密度 ρ 是*局部绝对可积的*是指对任意分图表 (U,α), 函数 ρ_α 是绝对可积的. 由第八章的最后面的那个命题知道, 这个定义仍旧与总图表的选取无关.

设 ρ 是 M 上的一个密度, x 为 M 的一个点. 这时谈及 ρ 在 x 的值是没有意义的. 然而 (3.1) 式表明谈及 ρ 在 x 的符号却是有意义的. 准确地说, 如果对 x 的一个分图表 (U,α) 有 $\rho_\alpha(\alpha(x)) > 0$, 则称在 x

$$\rho > 0. \tag{3.2}$$

等式 (3.1) 表明, 如果 $\rho_\alpha(\alpha(x)) > 0$, 则 $\rho_\beta(\beta(x)) > 0$ 对任何 x 的其他分图表都成立. 类似地, 说在 x 处 $\rho < 0$, 在 $x, \rho > 0$ 或在 $x, \rho \neq 0$ 都是有意义的.

定义 3.2 设 ρ 为 M 上一个密度. 记为 supp ρ 的 ρ 的支集是指 M 中那些使 ρ 不为零的点的集合的闭包. 就是说

$$\text{supp } \rho = \overline{\{x : 在 x, \rho \neq 0\}}.$$

设 ρ_1 和 ρ_2 为密度. 定义它们的和为对每个分图表 (U,α) 有

$$(\rho_1 + \rho_2)_\alpha = \rho_{1\alpha} + \rho_{2\alpha}. \tag{3.3}$$

由此立刻可知, (3.3) 式的右端满足转移法则 (3.1) 式, 故定义了 M 上的密度.

设 ρ 为密度, f 为函数, 我们定义密度 $f\rho$ 为

$$(f\rho)_\alpha = f_\alpha \rho_\alpha. \tag{3.4}$$

由于对函数的转移法则, 可以立即验证 (3.1) 成立.

明显地我们有

$$\text{supp } (\rho_1 + \rho_2) \subset \text{supp } \rho_1 \cup \text{supp } \rho_2 \tag{3.5}$$

以及

$$\text{supp } (f\rho) = \text{supp } f \cap \text{supp } \rho. \tag{3.6}$$

如果在 x 有 $\rho_2 - \rho_1 \geqslant 0$, 我们则记为在 x 有 $\rho_1 \leqslant \rho_2$, 而如果对所有的 $x \in M$ 有 $\rho_1 \leqslant \rho_2$, 则说 $\rho_1 \leqslant \rho_2$.

设 P 表示具有紧支集的局部绝对可积的密度组成的空间. 我们将 P 看作是个向量空间, 而当 f 是个 (有界的) 局部可度的函数同时 $\rho \in P$ 时则 $f\rho$ 属于 P.

定理 3.1 存在 P 上的惟一线性函数 $\int$, 满足下述条件: 如果 $\rho \in P$ 使得 supp $\rho \subset U$, 而 (U,α) 是 M 的一个分图表, 则

$$\int \rho = \int_{\alpha(U)} \rho_\alpha. \tag{3.7}$$

证明 我们首先证明最多存在一个满足 (3.7) 式的线性函数. 设 $\mathfrak{a}$ 为 M 的一个总图表, $\{g_i\}$ 为从属于 $\mathfrak{a}$ 的 1 的分解. 对每个 j 选取 $i(j)$ 使得

$$\text{supp } g_j \subset U_{i(j)}.$$

写出 $\rho = 1 \cdot \rho = \sum g_j \rho$. 因为 supp ρ 是紧的, 只有有限多个 $g_j\rho$ 项不恒为零. 因而此和式是有限的. 又由于 $\int$ 是线性的, 故

$$\int \rho = \int \sum g_j \rho = \sum \int g_j \rho.$$

由 (3.7) 式,

$$\int g_j\rho = \int_{\alpha_{i(j)}(U_{i(j)})} (g_j\rho)_{\alpha_{i(j)}}.$$

因此

$$\int \rho = \sum_j \int_{\alpha_{i(j)}(U_{i(j)})} (g_j\rho)_{\alpha_{i(j)}}. \tag{3.8}$$

因此 $\int$ 如果存在则必定由 (3.8) 式给出. 要建立 $\int$ 的存在性, 我们必须证明 (3.8) 式定义了一个 P 上的满足 (3.7) 式的线性函数. 线性性是显然的; 我们必须验证 (3.7) 式.

假设 $\operatorname{supp}\rho\subset U$, 其中 (U,α) 为某个分图表. 我们必须证明

$$\int_{\alpha(U)}\rho_\alpha = \sum_j \int_{\alpha_{i(j)}(U_{i(j)})} (g_j\rho)_{\alpha_{i(j)}}.$$

由于 $\rho=\sum g_j\rho$, 因此 $\rho_\alpha=\sum(g_j\rho)_\alpha$, 所以只要证明

$$\int_{\alpha(U)}(g_j\rho)_\alpha = \int_{\alpha_i(U_i)}(g_j\rho)_{\alpha_i} \tag{3.9}$$

即可, 其中 $\operatorname{supp} g_j\rho\subset U\cap U_i$. 由 (3.1) 式, 我们有

$$(g_j\rho)_\alpha = (g_j\rho)_{\alpha_i}\circ(\alpha_i\circ\alpha^{-1})\cdot|\det J_{\alpha_i\circ\alpha^{-1}}|,$$

故由 $\mathbb{R}^n$ 中的积分的变换定律知 (3.9) 式成立. □

由 (3.8) 式我们可以推出积分的许多有用的性质:

$$\text{如果 } \rho_1\leqslant\rho_2, \text{则} \int\rho_1\leqslant\rho_2. \tag{3.10}$$

事实上, 由于 $g_j\geqslant 0$, 故有 $(g_j\rho_1)_\alpha\leqslant(g_j\rho_2)_\alpha$ 对任意分图表 (U,α) 成立. 因而当我们运用 (3.8) 式时可由 $\mathbb{R}^n$ 中的相应事实推导出 (3.10) 式.

我们说集合 A 具有容度零是指如果 $A\subset A_1\cup\cdots\cup A_p$, 其中每个 A_i 为紧, 且对某个分图表 (U_i,α_i) 有 $A_\iota\subset U_i$, 使得 $\alpha_i(A_\iota)$ 在 $\mathbb{R}^n$ 中具有容度零. 容易看出任意有限个容度零的集合的并仍为容度零. 函数 e_A 明显是可度的.

我们说一个集合 $B\subset M$ 是可度的是指函数 e_B 是可度的. 对任何 $\rho\in P$, 定义 $\int_B\rho$ 为

$$\int_B\rho = \int e_B\rho. \tag{3.11}$$

于是由 (3.8) 式知, 如果 A 具容度零, 则对任何 $\rho \in P$ 有

$$\int_A \rho = 0.$$

因此我们在做积分时可以忽略掉容度为 0 的集合. 在实践中, 人们在计算积分时通常都利用这个性质而不是使用 (3.8) 式. 例如, 在计算 S^n 上一个积分时, 我们可以"忽略"任何一条经度线: 举例说, 如果

$$A = \{x \in S^n : x = (t, 0, \cdots, \pm\sqrt{1-t^2}) \in \mathbb{R}^{n+1}\},$$

则

$$\int_{S^n} \rho = \int_{S^n - A} \rho$$

对任意的 ρ 成立. 这意味着我们可以引进极坐标来计算 $\int_{S^n} \rho$(图 10.1) 并且用它们来表达 ρ. 因此在 S^2 中, 如果 $U = S^2 - A$, 而 α 是 U 上的极坐标分图表, 则

$$\int_{S^2} \rho = \int_0^{2\pi} \int_0^{\pi} \rho_\alpha d\theta d\varphi.$$

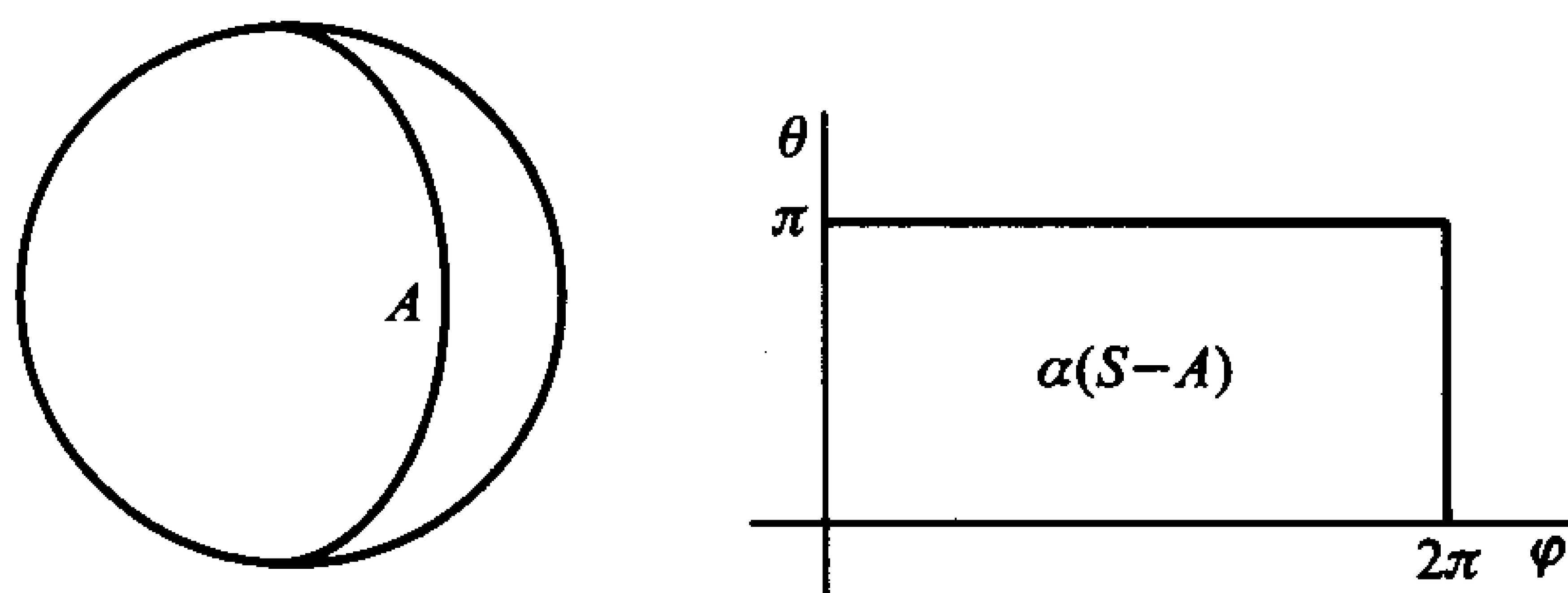


图 10.1

值得注意的是, 如果 N 是个维数小于 $\dim M$, 而 ψ 是 $N \to M$ 的一个可微映射, 则第八章的命题 7.3 表明如果 A 为 N 中任一紧子集, 则 $\psi(A)$ 在 M 中具有容度零. 在此意义下, 当我们在 M 上做积分时便可以忽略掉"低维的集合".

10.4 黎曼度量的体积密度

设M 为具黎曼度量 $\mathfrak{e}$ 的一个微分流形. 定义密度 $\sigma[= \sigma(\mathfrak{e})]$ 如下. 对每个有坐标 $x^1, \cdots, x^n$ 的分图表 (U, α), 设

$$\sigma_\alpha(\alpha(x)) = \left|\det\left[\left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x)\right)\right]\right|^{1/2} = |\det(g_{ij}(x))|^{1/2}. \qquad (4.1)$$

其中

$$\left[\left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x)\right)\right]$$

是一个矩阵, 其第 (ij) 元为向量

$$\frac{\partial}{\partial x^i}(x) \quad 和 \quad \frac{\partial}{\partial x^j}(x)$$

的内积, 故 (由第八章习题 8.1) 有

$\sigma_\alpha(\alpha(x))=$ 由在 $T_x(M)$ 上的欧几里得度量 $(\ ,\ )_{e,x}$ 下 $\partial/\partial x^i(x)$ 张成的平行六面体的体积.

容易看出 (4.1) 式实际上定义了一个密度. 设 (W,β) 为 x 的第二个分图表, 坐标为 $y^1,\cdots,y^n$. 于是

$$\frac{\partial}{\partial y^k}=\sum_i \frac{\partial x^i}{\partial y^k}\frac{\alpha}{\alpha x^i},$$

故

$$\sigma_\beta(\beta(x))=\left|\det\left[\left(\frac{\partial}{\partial y^k}(x), \frac{\partial}{\partial y^l}(x)\right)\right]\right|^{1/2}.$$

现对所有 k,l 有

$$\left(\frac{\partial}{\partial y^k}, \frac{\partial}{\partial y^l}\right)=\sum_{i,j}\frac{\partial x^i}{\partial y^k}\cdot\frac{\partial x^j}{\partial y^l}\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right).$$

我们写其为矩阵等式

$$\left[\left(\frac{\partial}{\partial y^k}, \frac{\partial}{\partial y^l}\right)\right]=\left[\frac{\partial x^i}{\partial y^k}\right]\left[\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)\right]\left[\frac{\partial x^j}{\partial y^l}\right],$$

从而

$$\begin{aligned}\sigma_\beta(\beta(x))&=\left|\det\left[\left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x)\right)\right]\det\left[\frac{\partial x^i}{\partial y^k}\right]\det\left[\frac{\partial x^j}{\partial y^l}\right]\right|^{1/2}\\&=\left|\det\left[\left(\frac{\partial}{\partial x^i}(x), \frac{\partial}{\partial x^j}(x)\right)\right]\right|^{1/2}\left|\det\left[\frac{\partial x^i}{\partial y^k}\right]\right|\\&=\sigma_\alpha(\alpha(x))\left|\det\left[\frac{\partial x^i}{\partial y^k}\right]\right|(x).\end{aligned}$$

如果 M 是具欧几里得度量的欧几里得空间中的开子集, 则对体积密度在任一可度集合上积分便产生出这个集合的通常欧几里得体积. 事实上, 如果

$x^1, \cdots, x^n$ 为对应于这个等同分图表的法正交坐标, 则 $g_{ij}(x)$ 在 $i \neq j$ 时为 0 而 $g_{ii} = 1$, 故 $\sigma_{\mathrm{id}} \equiv 1$, 因此

$$\int_A \sigma = \int_A 1 = \mu(A).$$

更一般地, 设 φ 为 k 维流形 M 到 $\mathbb{R}^n$ 的一个浸入, 使得 $\varphi(M)$ 是 $\mathbb{R}^n$ 中一个 k 维超平面中的开子集, 令 m 为 φ 在 M 上诱导的黎曼度量. 那么, 如果 σ 表示相应的体积密度, 则 $\int_A \sigma$ 是 $\varphi(A)$ 的 k 维欧几里得体积. 事实上, 经过一个欧几里得运动, 我们可设 φ 将 M 映入 $\mathbb{R}^k \subset \mathbb{R}^n$. 于是, 由于 φ 是个浸入, M 是 k 维的, 所以我们能用 $x^1, \cdots, x^k$ 作为 M 的坐标, 而且像前面那样, 得到由这些坐标的函数给出的 σ, 从而 $\int_A \sigma = \mu(\varphi(A))$.

现在设 φ_1 和 φ_2 为两个浸入: $M \to \mathbb{R}^n$. 设 (U, α) 为 M 上一个坐标分图表, 坐标为 $y^1, \cdots, y^k$. 如果 m_i 代表由 φ_i 诱导的黎曼度量, 则

$$\left(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\right)_{m_1} = \left(\frac{\partial \varphi_1}{\partial y^i}, \frac{\partial \varphi_1}{\partial y^j}\right)$$

和

$$\left(\frac{\partial}{\partial y^i}, \frac{\partial}{\partial y^j}\right)_{m_2} = \left(\frac{\partial \varphi_2}{\partial y^i}, \frac{\partial \varphi_2}{\partial y^j}\right),$$

其中在右端的内积是欧几里得内积. 设 σ_1 和 σ_2 为对应于 m_1 和 m_2 的体积密度. 于是

$$\sigma_{1\alpha} = \left|\det\left[\left(\frac{\partial \varphi_1}{\partial y^i}, \frac{\partial \varphi_1}{\partial y^j}\right)\right]\right|^{1/2}$$

和

$$\sigma_{2\alpha} = \left|\det\left[\left(\frac{\partial \varphi_2}{\partial y^i}, \frac{\partial \varphi_2}{\partial y^j}\right)\right]\right|^{1/2}.$$

特别, 给出一个 $L > 0$, 存在一个 $K = K(k, n, L)$ 使得如果对所有 $i = 1, \cdots, k$ 有

$$\left\|\frac{\partial \varphi_1}{\partial y^i}\right\| < L \quad 和 \quad \left\|\frac{\partial \varphi_2}{\partial y^i}\right\| < L$$

则由中值定理,

$$|\sigma_{1\alpha} - \sigma_{2\alpha}| \leqslant K\left(\left\|\frac{\partial \varphi_2}{\partial y^1} - \frac{\partial \varphi_1}{\partial y^1}\right\| + \cdots + \left\|\frac{\partial \varphi_2}{\partial y^k} - \frac{\partial \varphi_1}{\partial y^k}\right\|\right).$$

粗略地说, 这表明如果 φ_1 和 φ_2 靠近的意思是它们的导数是靠近的, 那么它们所诱导的密度也是靠近的.

我们把此评注用于下面的情形. 设 φ_1 为 M 到 $\mathbb{R}^n$ 的一个浸入, (W,α) 为 M 的一个分图表, 其坐标为 $y^1,\cdots,y^k$. 设 $U = W - C = \cup U_i$, 其中的 C 为某个容度 0 的闭集, 使得当 $l \neq l'$ 时 $U_l \cap U_{l'} = \varnothing$. 对每个 l 让 z_l 为 U_l 中一点, 其坐标为 $\prec y_l^1,\cdots,y_l^k \succ$, 并且对 $z = \prec y^1,\cdots,y^k \succ$ 定义 φ_2 为

$$\varphi_2(y^1,\cdots,y^k) = \varphi_1(z_l) + \sum (y^i - y_l^i)\frac{\partial \varphi_1}{\partial y^i}(z_l),$$

其中 $z \in U_l$(见图 10.2).

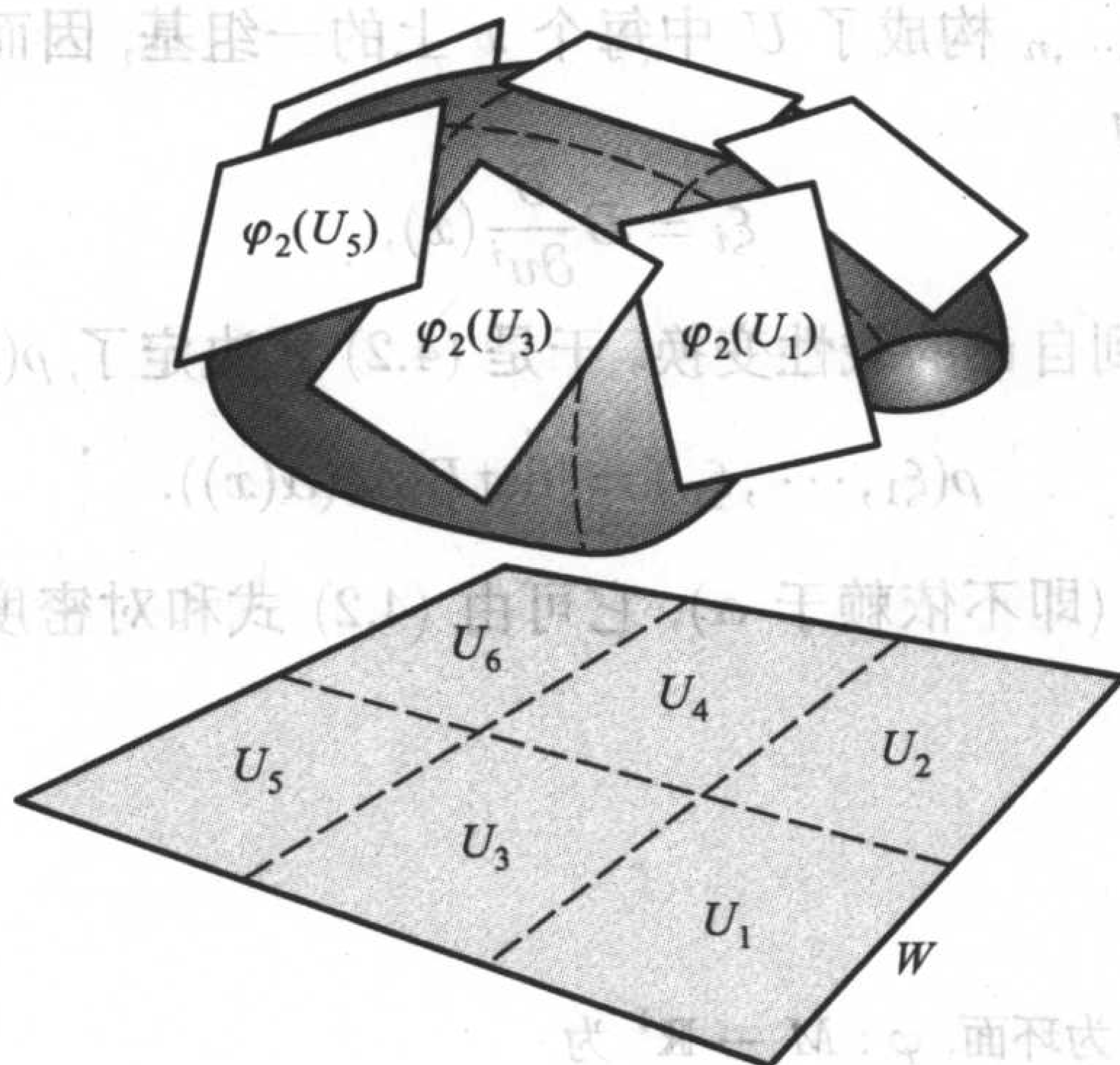


图 10.2

如果这些 U_l 充分小, 则

$$\left\| \frac{\partial \varphi_2}{\partial y^i} - \frac{\partial \varphi_1}{\partial y^i} \right\|$$

也是小的. 更一般地, 我们可选取 φ_2 为任何一个在每个 U_l 上逼近 φ_1 的仿射线性映射. 因而我们看到, *在由 φ 诱导的黎曼度量下的 W 的体积是逼近于 $\varphi(W)$ 的多面体(表面)的体积的极限*. 这里的逼近必须是在斜率的意义下的(即导数必须靠近) 而不仅仅是在所处位置的意义下.

体积密度的构造可以进行推广, 并提出了密度概念的另一种定义. 事实上设 ρ 是一个规则, 它对 M 中每个 x 指定一个在 $T_x(M)$ 中 n 个切向量上的函数 ρ_x, 从属于规则

$$\rho_x(A\xi_1,\cdots,A\xi_n) = |\det A|\rho_x(\xi_1,\cdots,\xi_n), \tag{4.2}$$

其中 $\xi_i \in T_x(M), A: T_x(M) \to T_x(M)$ 是个线性变换. 于是我们看到 ρ 决定了一个密度, 其定义是

$$\rho_\alpha(\alpha(x)) = \rho\left(\frac{\partial}{\partial u^1}(x), \cdots, \frac{\partial}{\partial u^n}(x)\right), \tag{4.3}$$

其中 (U, α) 是个具坐标 $u^1, \cdots, u^n$ 的分图表.(4.3) 式定义了一个密度的结论立即由 (4.2) 式和在坐标变换下 $\dfrac{\partial}{\partial u^i}$ 的变换定律得到.

反之, 以 ρ_α 方式给出一个密度 ρ, 由 (4.3) 式定义了 $\rho\left(\dfrac{\partial}{\partial u^1}, \cdots, \dfrac{\partial}{\partial u^n}\right)$. 因为向量 $\{\partial/\partial u^i\}_{i=1,\cdots,n}$ 构成了 U 中每个 x 上的一组基, 因而 $T_x(M)$ 中任何的 $\xi_1, \cdots, \xi_n$ 可以写为

$$\xi_i = B\frac{\partial}{\partial u^i}(x),$$

其中 B 是 $T_x(M)$ 到自己的线性变换. 于是 (4.2) 式决定了 $\rho(\xi_1, \cdots, \xi_n)$ 为

$$\rho(\xi_1, \cdots, \xi_n) = |\det B|\rho_\alpha(\alpha(x)). \tag{4.4}$$

这个定义是相容的 (即不依赖于 α), 它可由 (4.2) 式和对密度的变换定律 (3.1) 式推出.

习题

4.1 设 $M = S^1 \times S^1$ 为环面, $\varphi: M \to \mathbb{R}^4$ 为

$$x^1 \circ \varphi(\theta_1, \theta_2) = \cos\theta_1,$$
$$x^2 \circ \varphi(\theta_1, \theta_2) = \sin\theta_1,$$
$$x^3 \circ \varphi(\theta_1, \theta_2) = 2\cos\theta_2,$$
$$x^4 \circ \varphi(\theta_1, \theta_2) = 2\sin\theta_2,$$

其中 $x^1, \cdots, x^4$ 为 $\mathbb{R}^4$ 上的直角坐标, θ_1 和 θ_2 是 M 上的角坐标.

(a) 用坐标 θ_1, θ_2 表达由 φ(从 $\mathbb{R}^4$ 上的欧几里得度量) 在 M 上诱导黎曼度量.[就是说, 计算 $g_{ij}(\theta_1, \theta_2)$.]

(b) 关于这个黎曼度量, M 的体积是什么?

4.2 考虑由 $S^1 \times S^1$ 到 $\mathbb{E}^3$ 的浸入 φ 所诱导的黎曼度量. φ 的定义是

$$x \circ \varphi(u, v) = (a - \cos u)\cos v,$$
$$y \circ \varphi(u, v) = (a - \cos u)\sin v,$$
$$z \circ \varphi(u, v) = \sin u,$$

其中 u, v 为角坐标, $a > 2$. 在此度量下的 $S^1 \times S^1$ 的总曲面面积是什么?

4.3 设 φ 为由 xy 平面上一个区域 U 到 $\mathbb{E}^3$ 的映射, 其公式为

$$\varphi(x, y) = (x, y, F(x, y)),$$

使得 $\varphi(U)$ 为曲面 $z = F(x, y)$(见图 10.3). 证明此曲面面积由

$$\int_U \sqrt{1 + \left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2}$$

给出.

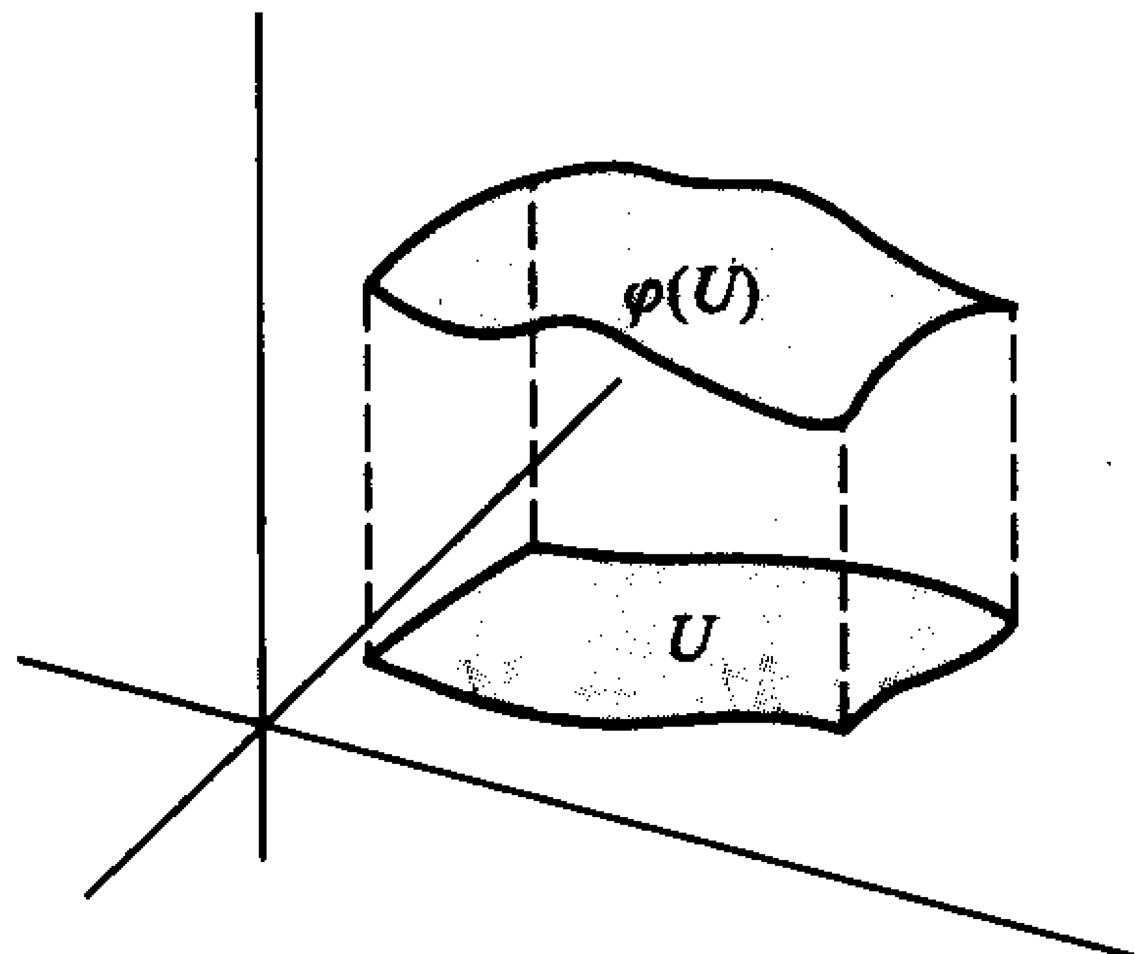


图 10.3

4.4 求抛物面

$$z = x^2 + y^2, \ x^2 + y^2 \leqslant 1$$

的面积.

4.5 设 $U \subset \mathbb{R}^2, \varphi : U \to \mathbb{E}^3$, 由

$$\varphi(u, v) = (x(u, v), y(u, v), z(u, v))$$

给出, 其中 x, y, z 为 $\mathbb{E}^3$ 的直角坐标. 证明曲面 $\varphi(U)$ 的面积为

$$\int_U \sqrt{\left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\frac{\partial z}{\partial v} - \frac{\partial y}{\partial v}\frac{\partial z}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\frac{\partial x}{\partial v} - \frac{\partial z}{\partial v}\frac{\partial x}{\partial u}\right)^2}.$$

4.6 计算 $\mathbb{E}^3$ 中单位球面的表面积.

4.7 设 M_1 和 M_2 是微分流形, σ 为 M_2 上的密度, 其处处不为零. 对于 $M_1 \times M_2$ 上的每个密度 ρ, 以及每个乘积分图表 $(U_1 \times U_2, \alpha_1 \times \alpha_2)$ 和每个 $x_2 \in U_2$, 定义函数 $\rho_{1\alpha_1}(\cdot, x_2)$ 为

$$\rho_{1\alpha_1}(v_1, x_2)\sigma_{\alpha_2}(\alpha_2(x_2)) = \rho_{\alpha_1 \times \alpha_2}(v_1, \alpha_2(x_2)),$$

其中 $v_1 \in \alpha_1(U_1)$.

(a) 证明 $\rho_{1\alpha_1}(v_1, x_2)$ 与分图表 (U_2, α_2) 的选取无关.

(b) 证明对每个固定的 $x_2 \in M_2$, 函数 $\rho_{1\alpha_1}(\cdot, x_2)$ 定义了 M_1 上的一个密度. 我们记此密度为 $\rho_1(x_2)$.

(c) 证明如果 ρ 是 $M_1 \times M_2$ 上具紧支集的光滑密度, σ 为光滑, 则 $\rho_1(x_2)$ 是 M_1 上一个具紧支集的光滑密度.

(d) 设 ρ 同于 c) 中的 ρ. 定义 M_2 上的函数 F_ρ 为

$$F_\rho(x_2) = \int_{M_1} \rho_1(x_2).$$

概述你会如何证明下述事实: F_ρ 是 M_2 上具紧支集的光滑函数, 并且

$$\int_{M_1 \times M_2} \rho = \int_{M_2} F_\rho \cdot \sigma.$$

10.5 密度的拉回和它的李导数

设 $\varphi : M_1 \to M_2$ 为微分同胚, ρ 为 M_2 上的一个密度. 定义 M_1 上的密度 $\varphi^*\rho$ 为

$$\varphi^*\rho(\xi_1, \cdots, \xi_n) = \rho(\varphi_*\xi_1, \cdots, \varphi_*\xi_n), \tag{5.1}$$

其中 $\xi_i \in T_x(M_1), \varphi_* = \varphi_{*x}$. 要证明 $\varphi^*\rho$ 实际上是一个密度, 我们就必须验证 (4.2) 式对 $T_x(M_1)$ 上任一个线性变换 A 成立. 然而

$$\begin{aligned}\varphi^*\rho(A\xi_1, \cdots, A\xi_n) &= \rho(\varphi_* A\xi_1, \cdots, \varphi_* A\xi_n)\\ &= \rho(\xi_* A\varphi_*^{-1}\varphi_*\xi_1, \cdots, \varphi_* A\varphi_*^{-1}\varphi_*\xi_n)\\ &= |\det\varphi_* A\varphi_*^{-1}|\rho(\varphi_*\xi_1, \cdots, \varphi_*\xi_n)\\ &= |\det A|\varphi^*\rho(\xi_1, \cdots, \xi_n),\end{aligned}$$

它就是所要的等式.

设 (U,α) 和 (W,β) 为 M_1 和 M_2 上相容的分图表, 各自的坐标为 $u^1, \cdots, u^n$ 和 $w^1, \cdots, w^n$. 于是对 U 中所有点我们由 (4.3) 式有

$$\begin{aligned}(\varphi^*\rho)_\alpha(\alpha(\cdot)) &= \rho\left(\varphi_*\frac{\partial}{\partial u^1}, \cdots, \varphi_*\frac{\partial}{\partial u^n}\right) = \left|\det\left(\frac{\partial w^j}{\partial u^i}\right)\right| \rho\left(\frac{\partial}{\partial w^1}, \cdots, \frac{\partial}{\partial w^n}\right)\\ &= \left|\det\left(\frac{\partial w^j}{\partial u^i}\right)\right| \rho_\beta(\beta \circ \varphi(\cdot)).\end{aligned}$$

换句话说, 我们有

$$(\varphi^*\rho)_\alpha = |\det J_{\beta\circ\varphi\circ\alpha^{-1}}|\rho(\beta \circ \varphi \circ \alpha^{-1}(\cdot)) \tag{5.2}$$

称密度 $\varphi^*\rho$ 为 ρ 在 φ^* 下的拉回. 明显地有

$$\varphi^*(\rho_1 + \rho_2) = \varphi^*(\rho_1) + \varphi^*(\rho_2)$$

以及对任意函数 f 有

$$\varphi^*(f\rho) = \varphi^*(f)\varphi^*(\rho).$$

直接从定义得到

$$\operatorname{supp} \varphi^* \rho = \varphi^{-1}[\operatorname{supp} \rho].$$

命题 5.1 设 $\varphi: M_1 \to M_2$ 为微分同胚, ρ 为 M_2 上具紧支集的局部绝对可积密度. 于是有

$$\int \varphi^* \rho = \int \rho. \tag{5.3}$$

证明 只要对

$$\operatorname{supp} \rho \subset \varphi(U)$$

的情形证明 (5.3) 式就可以了, 其中 (U, α) 为 M_1 的某个分图表, 且 $\varphi(U) \subset W$, 而 (W, β) 又是 M_2 的一个分图表. 事实上, 所有这类 $\varphi(U)$ 的集合是 M_2 的一组开覆盖, 从而我们可以选取一个从属于它的一个 1 的分解 $\{g_i\}$. 如果我们写成 $\rho = \sum g_j \rho$, 则此和有限, 而且每个 $g_j \rho$ 有上面看到的性质. 由于 (5.3) 式的两端为线性, 我们便知道只须对每个项去证明 (5.3) 式即可.

现在如果 $\operatorname{supp} \rho \subset \varphi(U)$, 则

$$\int \rho = \int_{\beta(W)} \rho_\beta = \int_{\beta \circ \varphi(U)} \rho_\beta$$

和

$$\begin{aligned}\int \varphi^* \rho = \int_{\alpha(U)} (\varphi^* \rho)_\alpha &= \int_{\alpha(U)} \rho_\beta (\beta \circ \varphi \circ \alpha^{-1}) |\det J_{\beta \circ \varphi \circ \alpha^{-1}}| \\ &= \int_{\beta \circ \varphi(U)} \rho_\beta,\end{aligned}$$

从而 (5.3) 式成立. □

现设 φ_t 为 M 上的一个单参数群, 其无穷小生成元为 X, 设 ρ 为 M 上一个密度, (U, α) 为分图表, W 为 U 的一个开子集使得 $\varphi_t(W) \subset U$ 对 $|t| < \varepsilon$ 的所有 t 成立. 于是对 $v \in \alpha(W)$ 有

$$(\varphi_t^* \rho)_\alpha(v) = \rho_\alpha(\Phi_\alpha(v, t)) \left| \det \left(\frac{\partial \Phi_\alpha}{\partial v} \right)_{(v,t)} \right|,$$

其中的 $\Phi_\alpha(v, t) = \alpha \circ \varphi_t \circ \alpha^{-1}(v)$ 而 $(\partial \Phi_\alpha / \partial v)_{(v,t)}$ 是 $v \mapsto \Phi_\alpha(v, t)$ 的雅可比. 我们想要计算这个表达式关于 t 的在 $t = 0$ 的导数. 现在 $\Phi_\alpha(v, 0) = v$, 从而

$$\det \left(\frac{\partial \Phi_\alpha}{\partial v} \right)_{(v,0)} = 1.$$

因此, 我们得到对于靠近 0 的 t 有

$$\det\left(\frac{\partial \Phi_\alpha}{\partial v}\right)_{(v,t)} > 0.$$

因此我们可以略去绝对值符号, 并写成

$$\left.\frac{d(\varphi_t^*\rho)_\alpha}{dt}\right|_{t=0} = \left.\frac{d\rho_\alpha(\Phi_\alpha)}{dt}\right|_{t=0} + \rho_\alpha(v)\left.\frac{d}{dt}\left(\det\frac{\partial \Phi_\alpha}{\partial v}\right)\right|_{t=0}.$$

我们只要简单地用链规则去计算右端的第一个导数, 得到

$$d\rho_\alpha\left(\frac{\partial \Phi_\alpha}{\partial t}\right) = d\rho_\alpha(X_\alpha(v)).$$

用坐标 $x^1,\cdots,x^n$ 表示则可写为

$$\frac{d\rho_\alpha(\Phi_\alpha(v,t))}{dt} = \sum \frac{\partial \rho_\alpha}{\partial x^i}X_\alpha^i,$$

其中 $X_\alpha = \prec x_\alpha^1,\cdots,X_\alpha^n \succ$.

要计算右端的第二项我们需要做一点预先的观察. 设 $A(t)=(a_{ij}(t))$ 为 t 的可微的矩阵值函数, 满足 $A(0)=\mathrm{id}=(\delta_i^j)$. 于是

$$\frac{d(\det A(t))}{dt} = \lim \frac{1}{t}(\det A(t)-1).$$

现 $a_{ii}(0)=0, a_{ij}(0)=0 (i\neq j)$. 称 A 是可微的是说每个函数 $a_{ij}(t)$ 是可微的. 因而我们可以找到一个常数 K 使得 $i\neq j$ 时 $|a_{ij}(t)|\leqslant K$ 而 $|a_{ii}(t)-1|\leqslant K|t|$. 在 $\det A(t)$ 的展开式中那些不以 t^2 趋于零的惟一项是对角线的乘积 $a_{11}(t)\cdots a_{nn}(t)$. 事实上, 任何 $\sum \pm a_{1i_1}(t)\cdots a_{ni_n}(t)$ 中的其他项至少包含了两个不在对角线上的因子, 因而至少以 t^2 趋向零. 于是

$$\begin{aligned}\frac{d}{dt}(\det A(t)) &= \lim\nolimits_{t\to 0}\frac{1}{t}(a_{11}(t)\cdots a_{nn}(t)-1)\\ &= a_{11}'(0)+\cdots+a'_{nn}(0)\\ &= \mathrm{tr}A'(0).\end{aligned}$$

如果我们取 $A=\partial\Phi_\alpha/\partial v$, 我们得到

$$\frac{d}{dt}\left(\det\frac{\partial \Phi_\alpha}{\partial v}\right) = \mathrm{tr}\frac{\partial X_\alpha}{\partial v} = \sum \frac{\partial X_\alpha^i}{\partial x^i}$$

因此

$$\frac{d(\varphi_t^*\rho)_\alpha}{dt}=\sum\frac{\partial\rho_\alpha}{\partial x^i}X_\alpha^i+\rho_\alpha\frac{\partial X_\alpha^i}{\partial x^i}=\sum\frac{\partial}{\partial x^i}(\rho_\alpha X_\alpha^i).$$

重写出来就是

命题 5.2 设 φ_t 为 M 的微分同胚的单参数群, 其无穷小生成元为 X, ρ 为 M 上的一个可微密度. 于是

$$D_X\rho=\lim_{t\to 0}\frac{\varphi_t^*\rho-\rho}{t}$$

存在, 并且局部地由

$$(D_X\rho)_\alpha=\sum\frac{\partial(\rho_\beta X_\alpha^i)}{\partial x^i}$$

给出, 其中在分图表 (U,α) 上 $X=\prec X_\alpha^1,\cdots,X_\alpha^n\succ$.

有时称密度 $D_X\rho$ 为 (X,ρ) 的散度, 以 $\mathrm{div}\prec X,\rho\succ$ 表示. 因此 $\mathrm{div}\prec X,\rho\succ=D_X\rho$ 是个密度, 它在 (U,α) 上由

$$(\mathrm{div}\prec X,\rho\succ)_\alpha=\sum\frac{\partial}{\partial x^i}(X_\alpha^i\rho_\alpha)$$

给出.

现设 ρ 为一可微密度, A 为紧的可度集合, 则

$$\begin{aligned}\int_{\varphi_t(A)}\rho&=\int_M e_{\varphi_t(A)}\rho=\int_M\varphi_t^*(e_{\varphi_t(A)}\rho)\\&=\int(\varphi_t^*e_{\varphi_t(A)})(\varphi_t^*\rho)=\int e_A\varphi_t^*(\rho)\\&=\int_A\varphi_t^*\rho.\end{aligned}$$

因此

$$\frac{1}{t}\left(\int_{\varphi_t(A)}\rho-\int_A\rho\right)=\int_A\frac{1}{t}(\varphi_t^*\rho-\rho).$$

应用 1 的分解, 我们可以容易看出在积分号内的极限是一致的, 从而有公式

$$\frac{d}{dt}\left(\int_{\varphi_t(A)}\rho\right)\bigg|_{t=0}=\int_A D_X\rho=\int_A \operatorname{div}\prec X,\rho\succ .$$

10.6 散度定理

设 φ 为微分流形 M 上的一个流, 其无穷小生成元为 X. 设 ρ 为属于 P 的一个密度, A 为 M 的一个可度子集. 于是对小的 t 值, 我们预料差值 $\int_{\varphi_t(A)}\rho-\int_A\rho$ 只依赖于在 A 的边界所发生的情况 (图 10.4). 在取极限中, 我们预料 $\int_{\varphi_t(A)}\rho$ 在 $t=0$ 的导数 (它由 $\int_A \operatorname{div}\prec X,\rho\succ$ 给出) 会由在 ∂A 上的某个积分给出. 为了详细阐述这样的结果, 我们首先必须挑出一类集合, 使其边缘足够好到让我们能够在它们上面进行积分. 于是我们给出下面的定义.

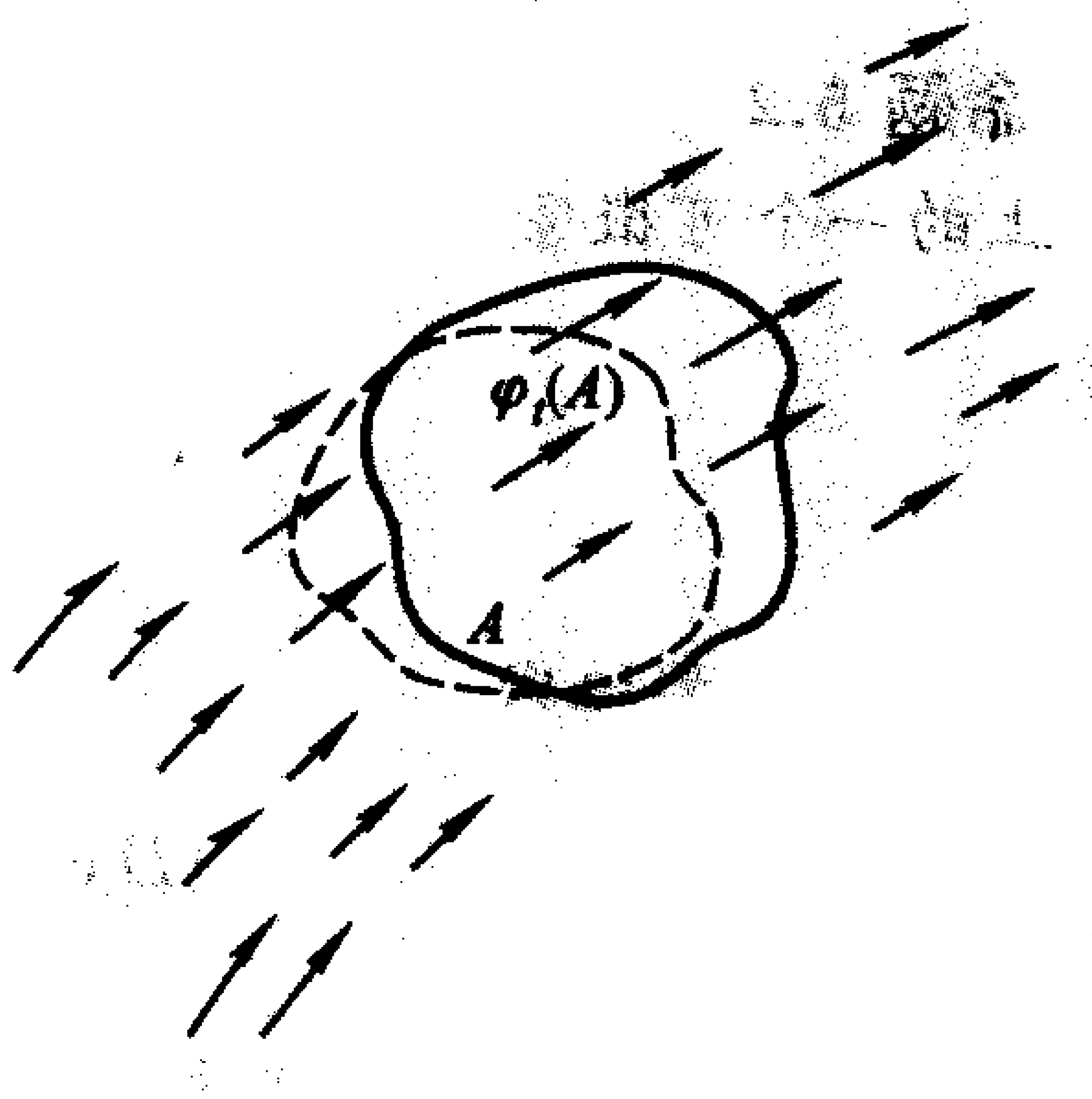


图 10.4

定义 6.1 设 M 为微分流形, D 为 M 的一个子集合. 我们称 D 是个具正则边界的区域是说对每点 $x\in M$ 存在一个 x 的分图表, 其坐标为 $x_\alpha^1,\cdots,x_\alpha^n$, 使得下面的三种可能性中的一个成立:

(i) $U\cap D=\varnothing$;

(ii) $U\subset D$;

(iii) $\alpha(U\cap D)=\alpha(U)\cap\{\boldsymbol{v}=\prec v^1,\cdots,v^n\succ\in\mathbb{R}^n : v^n\geqslant 0\}$.

注意, 如果 $x\notin\overline{D}$, 我们总可以找到一个 $(U,\alpha), x\in U$ 使得 (i) 成立. 如果 $x\in\operatorname{int}D$, 我们则总可找到一个 x 的分图表 (U,α) 使得 (ii) 成立. 这个强加于 D 上内容是不受限制的. 关键的条件是加在当 $x\in\partial D$ 的情形. 这时我们找不到 x 的分图表, 它满足 (i) 或者 (ii). 在这种情形, (iii) 表明 $\alpha(U\cap\partial D)$ 是 $\mathbb{R}^{n-1}$ 中的开子集 (图 10.5). 事实上, $\partial(U\cap\partial D)=\{\boldsymbol{v}\in\partial(U): v^n=0\}=\partial(U)\cap\mathbb{R}^{n-1}$, 其中我们将 $\mathbb{R}^{n-1}$ 看为 $\mathbb{R}^n$ 中那些最后一个分量为 0 的向量组成的子空间.

设 $\mathfrak{a}$ 为 M 上一个总图表, 使得 $\mathfrak{a}$ 的每个分图表满足 (i),(ii) 或 (iii) 中的一个. 对每个 $(U,\alpha)\in\mathfrak{a}$ 考虑映射 $\alpha\restriction\partial D: U\cap\partial D\longrightarrow\mathbb{R}^{n-1}\subset\mathbb{R}^n$[自然, 映射 $\alpha\restriction\partial D$

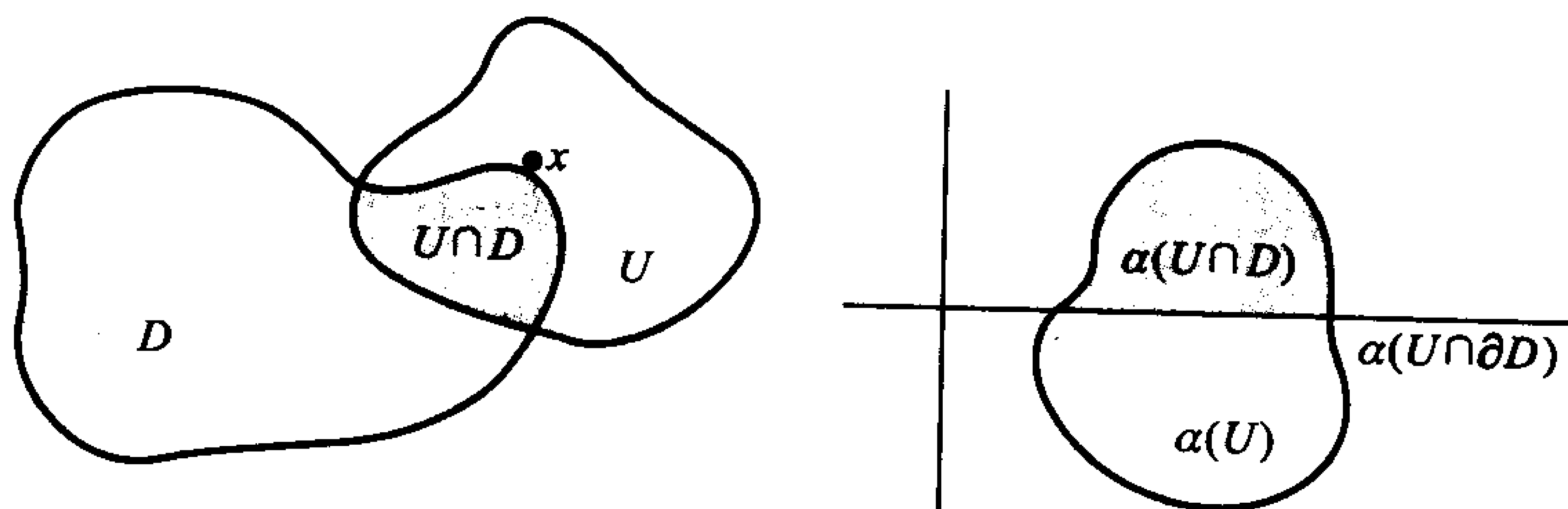


图 10.5

只对 (iii) 型的分图表具有非空的定义域.] 我们断言 $\{(U\cap\partial D, \alpha\restriction\partial D)\}$ 是 ∂D 上的一个总图表. 事实上, 设 (U,α), (W,β) 为 $\mathfrak{a}$ 中两个分图表使得 $U\cap W\cap\partial D\neq\varnothing$. 设 $x^1,\cdots,x^n$ 为 (U,α) 的坐标, 而 $y^1,\cdots,y^n$ 为 (W,β) 的. 映射 $\beta\circ\alpha^{-1}$ 由

$$\prec x^1,\cdots,x^n\succ\longmapsto\prec y^1(x^1,\cdots,x^n),\cdots,y^n(x^1,\cdots,x^n)\succ$$

给出. 在 $\alpha(U\cap W\cap\partial D)$ 上, 我们有 $x^n=0$ 和 $y^n=0$. 特别,

$$y^n(x^1,\cdots,x^{n-1},0)\equiv 0,$$

而函数 $y^1(x^1,\cdots,x^{n-1},0),\cdots,y^{n-1}(x^1,\cdots,x^{n-1},0)$ 可微. 这表明 $(\beta\restriction\partial D)\circ(\alpha\restriction\partial D)^{-1}$ 在 $\alpha(U\cap\partial D)$ 上可微. 从而我们得到了 ∂D 上的一个流形结构.

容易看出这个流形结构与我们所选取的这个特定的总图表无关. 我们以 ι 表示映射 $\partial D\longrightarrow M$, 它把每个点 $x\in\partial D$ 看作 M 中一个元素, 即映到自己. 明显地 ι 是个可微映射.(事实上, $(U\cap\partial D,\alpha\restriction\partial D)$ 和 (U,α) 相对于映射 $\alpha\circ\iota\circ(\alpha\restriction\partial D)^{-1}$ 是相容的分图表, 而这个映射就是 $\mathbb{R}^{n-1}\longrightarrow\mathbb{R}^n$ 的那个映射.)

设 x 为 ∂D 的点, 并视其为 M 中的一个点, 又设 ξ 为 $T_x(M)$ 中的一个元素. 我们说 ξ 指向 D 内指的是对每条有 $C'(0)=\xi$ 的曲线 C, 对充分小正的 t 我们有 $C(t)\in D$(图 10.6). 在类型 (iii) 的分图表 (U,α) 下, 设 $\xi_\alpha=\prec\xi^1,\cdots,\xi^n\succ$. 那么, 明显地; ξ 指向 D 内当且仅当 $\xi^n>0$. 相似地, 一个切向量 ξ 指向 D 外 (显然的定义) 当且仅当 $\xi^n<0$. 如果 $\xi^n=0$, 则 ξ 切于边界, 即它在 $\iota_*T_x(\partial D)$ 中.

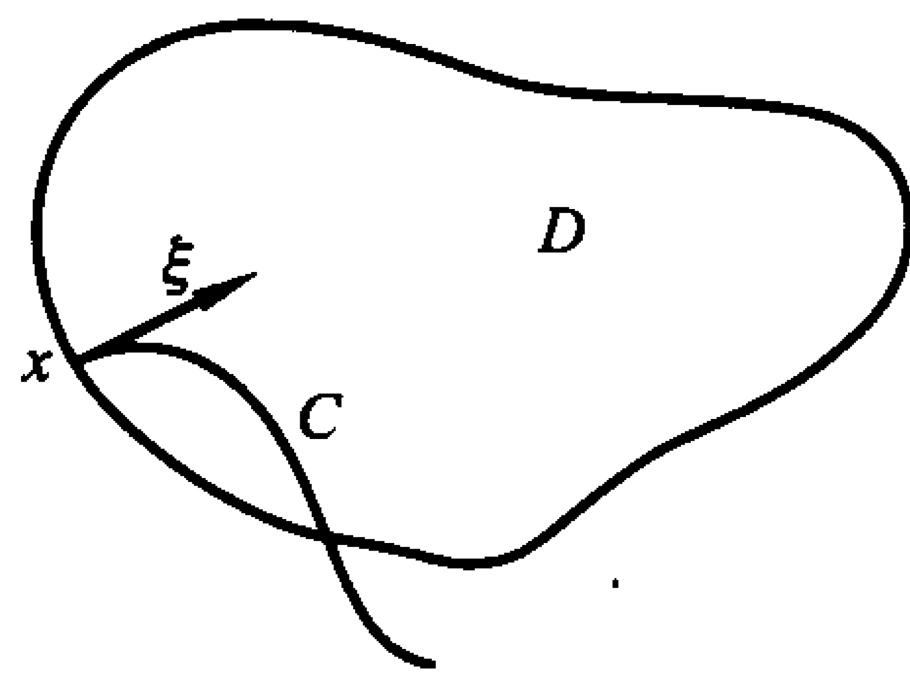


图 10.6

设 ρ 是 M 上一个密度, X 为 M 上的向量场.

定义 ∂D 上的密度 ρ_X 对于 $\xi_i \in T_x(\partial D)$ 的值为

$$\rho_X(\xi_1, \cdots, \xi_{n-1}) = \rho(\iota_*\xi_1, \cdots, \iota_*\xi_{n-1}, X(x)). \tag{6.1}$$

容易验证 (6.1) 式定义了一个密度.(留给读者作为练习.) 如果 (U,α) 是个型 (iii) 的分图表, $x \in U$, 而 $X_\alpha = \prec X^1, \cdots, X^n \succ$, 那么, 将 (4.3) 式用于分图表 $(U \cap \partial D, \alpha \restriction \partial D)$ 和密度 ρ_X, 我们看到

$$(\rho_X)_{\alpha \restriction \partial D} = \rho\left(\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^{n-1}}, X\right).$$

设 A 是 $T_x(M)$ 的一个线性变换, 定义为

$$A\frac{\partial}{\partial x^1} = \frac{\partial}{\partial x^1}, \cdots, A\frac{\partial}{\partial x^{n-1}} = \frac{\partial}{\partial x^{n-1}}, A\frac{\partial}{\partial x^n} = X.$$

那么 A 的矩阵是

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & X^1 \\ 0 & 1 & 0 & \cdots & 0 & X^2 \\ & & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & 1 & X^{n-1} \\ 0 & 0 & 0 & \cdots & 0 & X^n \end{bmatrix}$$

因此 $|\det A| = |X^n|$. 于是我们在 $\alpha(U \cap \partial D)$ 的所有点上有

$$(\rho_X)_{\alpha \restriction \partial D} = |X^n| \rho_\alpha. \tag{6.2}$$

现在可以叙述我们的结果了.

定理 6.1 (散度定理)[3] 设 D 为具正则边界的一个区域, $\rho \in P, X$ 为 M 上一个光滑向量场. 定义在 ∂D 上的函数 ε_X 为

$$\varepsilon_X(x) = \begin{cases} 1, & \text{如果 } X(x) \text{ 指向 } D \text{ 外}, \\ 0, & \text{如果 } X(x) \text{ 切于 } \partial D, \\ -1, & \text{如果 } X(x) \text{ 指向 } D \text{ 内}. \end{cases}$$

于是

$$\int_D \operatorname{div} \prec X, \rho \succ = \int_{\partial D} \varepsilon_X \rho_X. \tag{6.3}$$

[3]这个散度定理的公式化叙述和说明, 是 Richard Rasala 向我们建议的.

注 用型 (iii) 的一个分图表, 函数 ε_X 由

$$\varepsilon_X = -\mathrm{sgn} X^n \tag{6.4}$$

给出.

证明 设 $\mathfrak{a}$ 为 M 上一个总图表, 它的每个分图表属于前面所说的那三种类型之一. 设 $\{g_i\}$ 为从属于 $\mathfrak{a}$ 的 1 的分解. 记 $\rho = \sum g_i\rho$. 这是一个有限和. 由于 (6.3) 式两端对 ρ 都是线性的, 故只要对每个和项 $g_i\rho$ 进行验证即可. 改变记号 (由 ρ 代替 $g_i\rho$), 我们把问题化成在附加假定 $\mathrm{supp}\rho \subset U$ 去证明 (6.3) 式, 其中的 (U,α) 是类型 (i),(ii) 和 (iii) 中之一. 因此有三种情形要考虑.

情形 I $\mathrm{supp}\,\rho \subset U, U\cap \overline{D} = \varnothing$(见图 10.7). 于是 (6.3) 式的两边都为零, (6.3) 式正确.

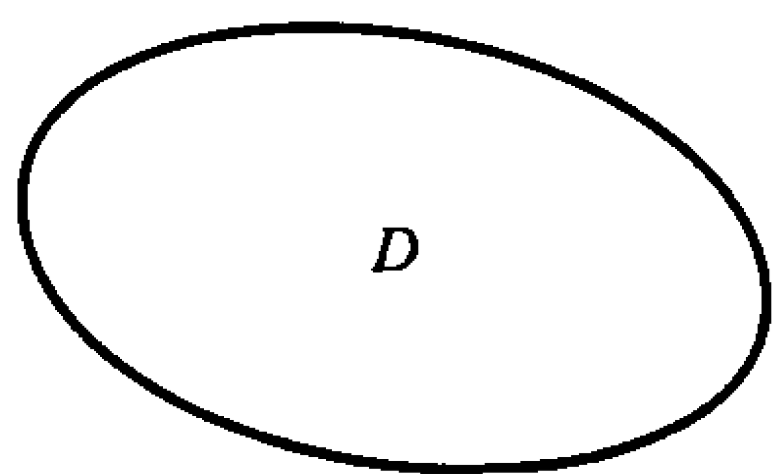


图 10.7

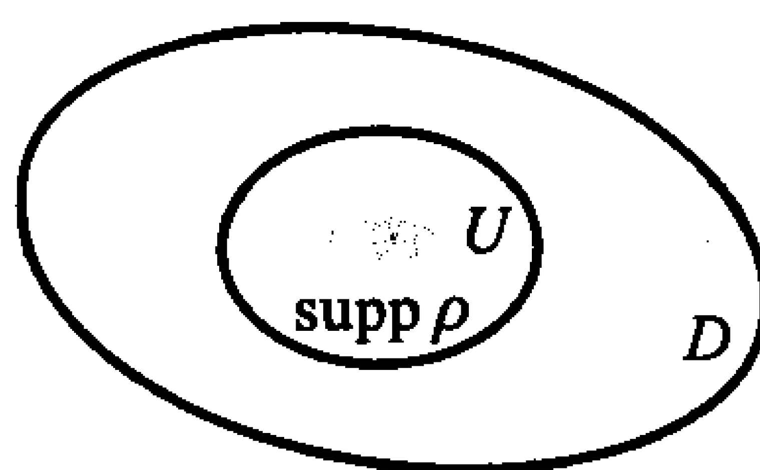


图 10.8

情形 II $\mathrm{supp}\rho \subset U, U\subset \mathrm{int}D$(见图 10.8). 于是 (6.3) 式的右端为 0, 我们必须证明左端也为零. 然而

$$\int_D \mathrm{div} \prec X,\rho \succ = \int_U \mathrm{div} \prec X,\rho \succ = \int_{\alpha(U)} \sum \frac{\partial(X^i\rho_\alpha)}{\partial x^i} = \sum \int_{\alpha(U)} \frac{\partial(X^i\rho_\alpha)}{\partial x^i}.$$

现在的每个函数 $X^i\rho_\alpha$ 的支集都在 $\alpha(U)$ 里面. 选取某个大的 R, 使得 $\alpha(U)\subset \square_{-R}^{R}$. 我们可以用 $\int_{\square_{-R}^{R}}$ 代替 $\int_{\alpha U}$. 把它在 $\alpha(U)$ 外的值取 0 来扩大 $X^i\rho_\alpha$ 的定义域到整个 $\mathbb{R}^n$(见图 10.9). 将此积分写成累次积分并首先关于 x^i 进行积分得到

$$\int_{\alpha(U)} \frac{\partial(X^i\rho_\alpha)}{\partial x^i} = \int X^i\rho_\alpha(\cdots,R,\cdots) \\ -X^i\rho_\alpha(\cdots,-R,\cdots)dx^1dx^2\cdots dx^{i-1}dx^i\cdots dx^n = 0.$$

这是因为函数 $X^i\rho_\alpha$ 在 $\alpha(U)$ 外为 0, 故上面这个积分为 0.

情形 III $\mathrm{supp}\rho$ 包含在型 (iii) 的一个分图表中 (见图 10.10). 于是

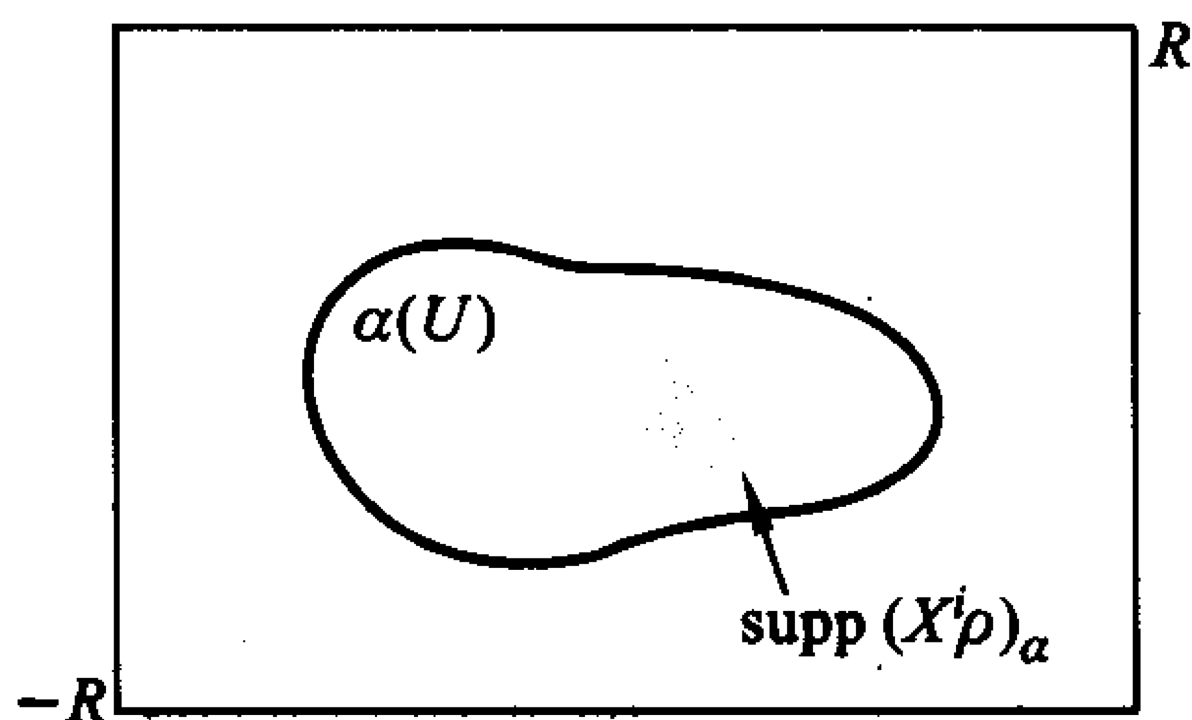


图 10.9

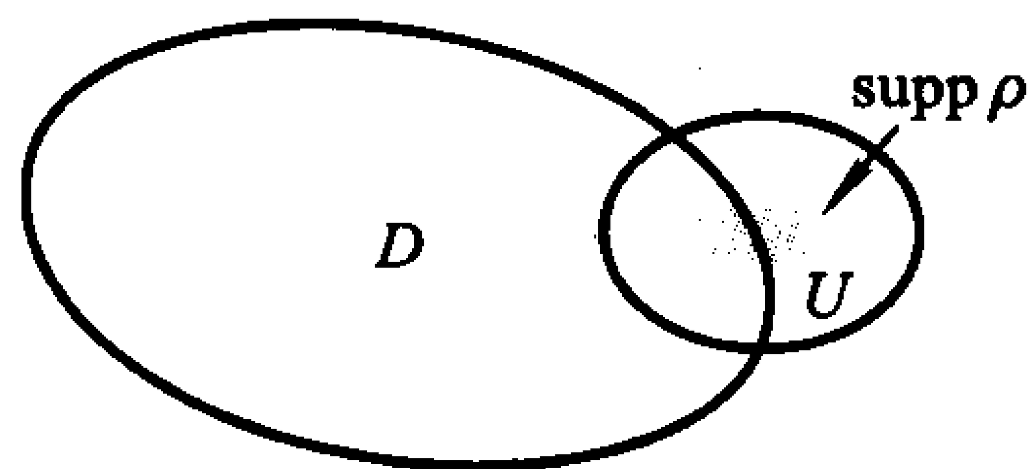


图 10.10

$$\int_D \operatorname{div} \prec X, \rho \succ = \int_{D\cap U} \operatorname{div} \prec X, \rho \succ = \sum \int_{\alpha(D\cap U)} \frac{\partial X^i \rho_\alpha}{\partial x^i}.$$

现在

$$\alpha(U \cap D) = \alpha(U) \cap \{\boldsymbol{v} : v^n \geqslant 0\}.$$

我们因此可以将积分区域换成

$$\square_{\prec -R,\cdots,-R,0\succ}^{\prec R,\cdots,R\succ}.$$

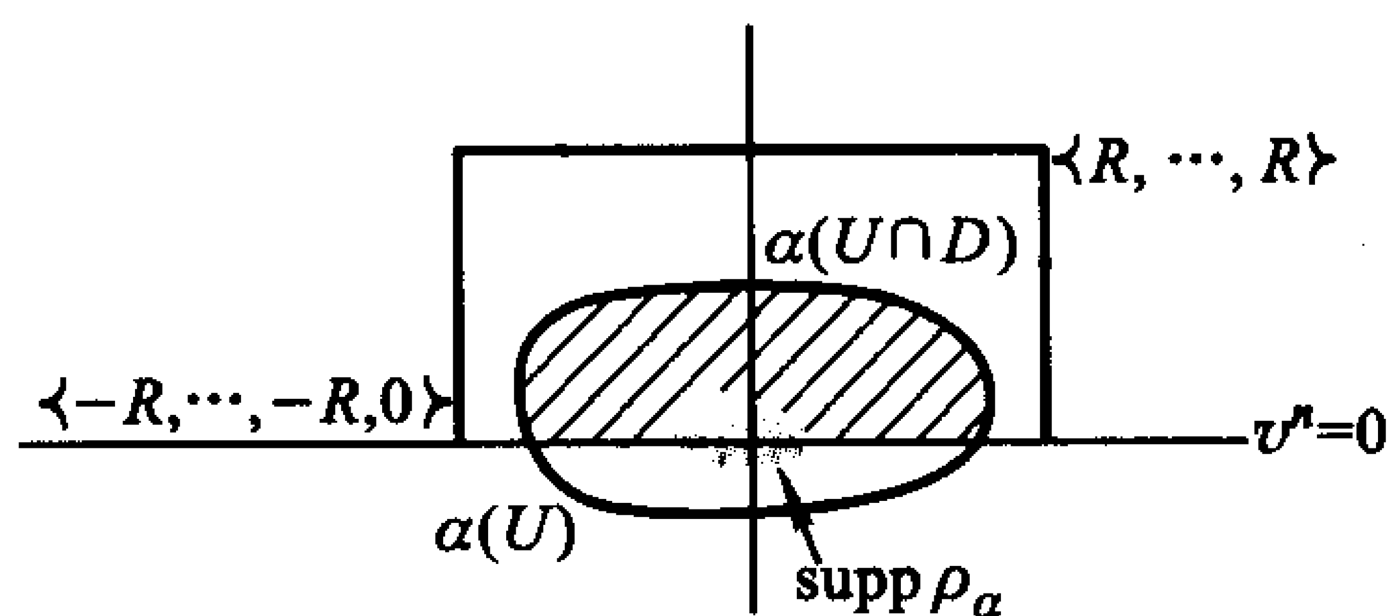


图 10.11

(见图 10.11) 对所有 $1 \leqslant i < n$, 在此和式中所有的积分如同上面一样为 0. 对 $i = n$ 我们有

$$\int_D \operatorname{div} \prec X, \rho \succ = -\int_{\mathbb{R}^{n-1}} X^n \rho_\alpha.$$

如果我们将此与 (6.2) 和 (6.4) 式比较, 这完全就是 (6.3) 式的断言. □

如果在流形 M 上给出了一个黎曼度量, 那么我们可以给出另一种版本的散度定理. 设 dV 是此黎曼度量下的体积密度, 从而对 $\xi_i \in T_x(M)$ 有

$$dV(\xi_1, \cdots, \xi_n) = |\det((\xi_i, \xi_j))|^{\frac{1}{2}}$$

是在切空间中 ξ_i 张成的平行六面体的体积 (关于由在切空间上的内积给出的欧几里得度量).

现在映射 ι 是个浸入, 因此我们在 ∂D 上得到一个诱导的黎曼度量. 设 dS 为 ∂D 上相对应的体积密度. 因此如果 $\{\xi_i\}_{i=1,\cdots,n-1}$ 为 $T_x(\partial D)$ 中的 $n-1$ 个向

量, $dS(\xi_1,\cdots,\xi_{n-1})$ 是由 $\iota_*T_x(\partial D)\subset T_x(M)$ 中 $\iota_*\xi_1,\cdots,\iota_*\xi_{n-1}$ 张成的平行六面体的 $(n-1)$ 维体积. 对任意 $x\in\partial D$ 设 $\boldsymbol{n}\in T_x(M)$ 为具单位长度的垂直于 $\iota_*T_x(\partial D)$ 的向量, 它指向 D 外 (图 10.12).

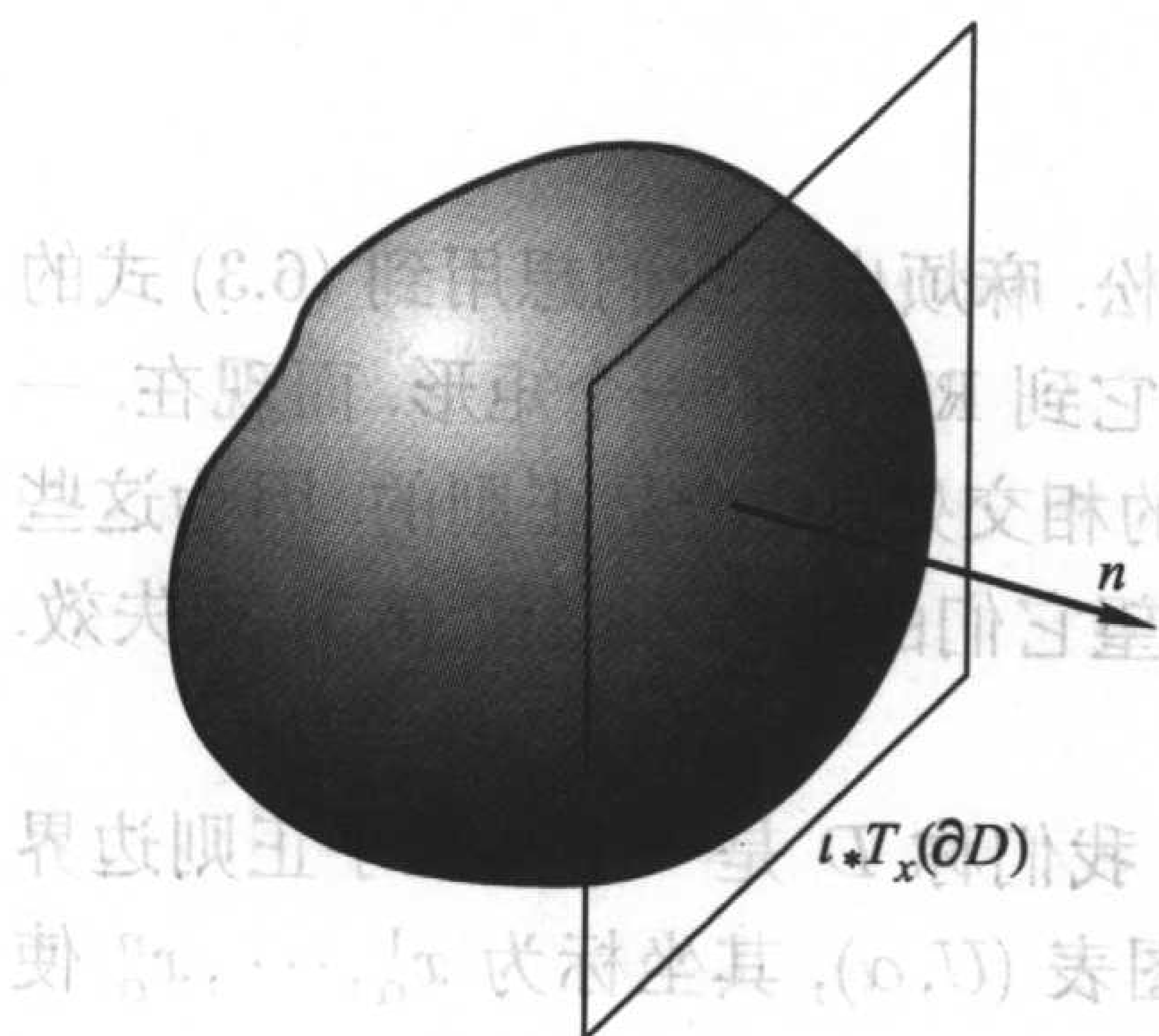


图 10.12

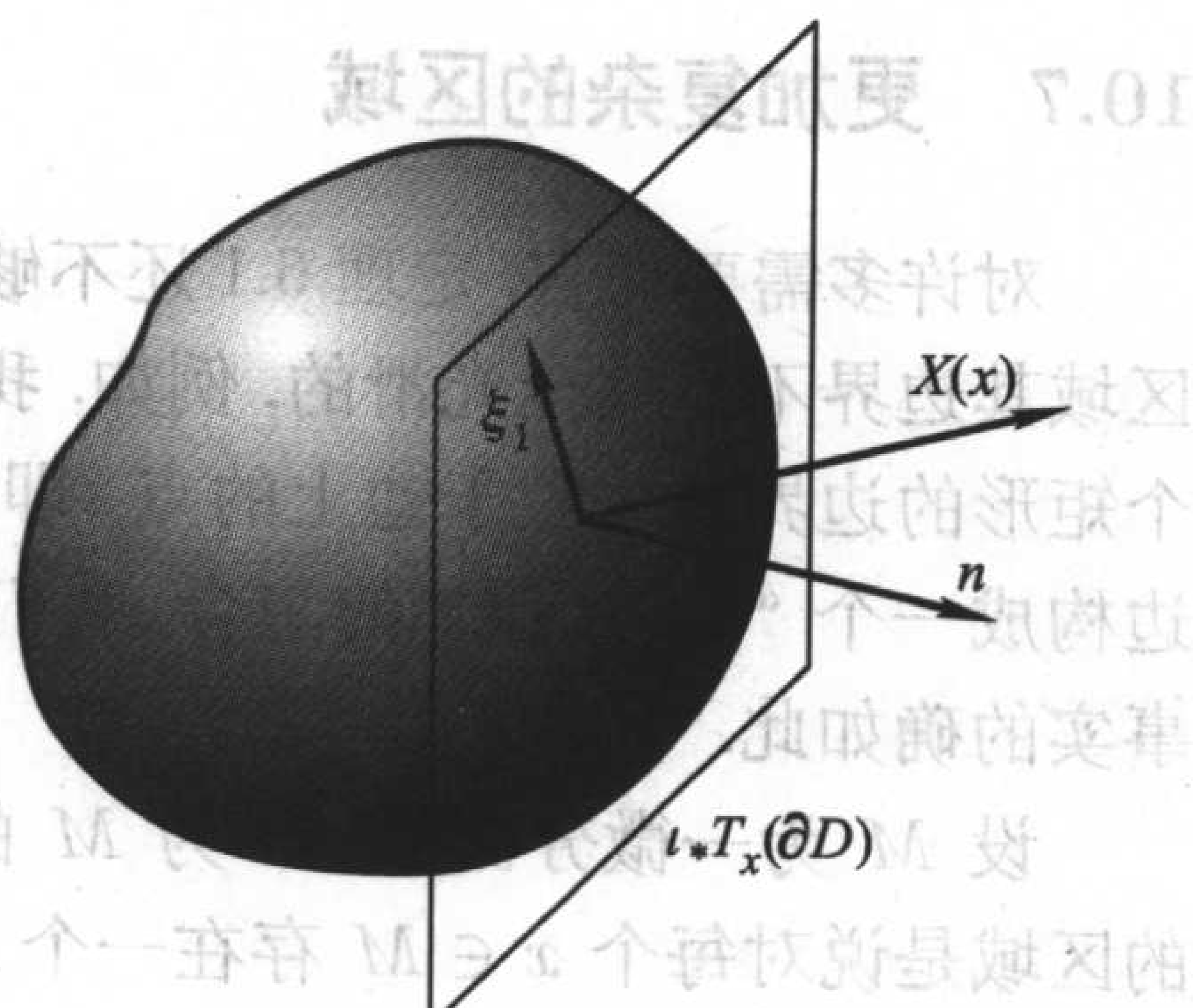


图 10.13

我们明显地有

$$dS(\xi_1,\cdots,\xi_{n-1})=dV(\iota_*\xi_1,\cdots,\iota_*\xi_{n-1},\boldsymbol{n}).$$

对任意向量 $X(x)\in T_x(M)$(图 10.13) 由 $\xi_1,\cdots,\xi_n,X(x)$ 张成的平行六面体的体积是 $|(X(x),\boldsymbol{n})|\,dS(\xi_1,\cdots,\xi_{n-1})$.[事实上, 写成

$$X(x)=(X(x),\boldsymbol{n})\boldsymbol{n}+\boldsymbol{m},$$

其中 $\boldsymbol{m}\in\iota_*T(\partial D)$.] 如果将此与 (6.1) 式作比较, 我们看到

$$dV_X=|(X(x),\boldsymbol{n})|\,dS. \tag{6.5}$$

另外, 显见有

$$\varepsilon(x)=\operatorname{sgn}(X(x),\boldsymbol{n}).$$

设 ρ 为 M 上的任意密度. 于是我们可写为

$$\rho=fdV,$$

其中 f 是一个函数. 又, 我们明显地有 $\rho_X=fdV_X$ 并且

$$\operatorname{div}\prec X,\rho\succ=\operatorname{div}\prec X,fdV\succ.$$

我们于是可重写 (6.3) 式为

$$\int_D \operatorname{div} \prec X, fdV \succ = \int_{\partial D} f \cdot (X, \boldsymbol{n}) dS \tag{6.6}$$

10.7 更加复杂的区域

对许多需要而言，定理 6.1 还不够充分宽松. 麻烦出在我们想用到 (6.3) 式的区域其边界不是完全光滑的. 例如，我们想用它到 $\mathbb{R}^n$ 中的一个矩形. 而现在，一个矩形的边界除去位于边上的点 (即两个面的相交处) 外全是正则的. 因为这些边构成一个 "$n-2$ 维" 的集合，我们可以期望它们的出现不会使 (6.3) 式失效. 事实的确如此.

设 M 为一微分流形，D 为 M 的子集. 我们称 D 是一个具几乎正则边界的区域是说对每个 $x \in M$ 存在一个 x 的分图表 (U, α)，其坐标为 $x_\alpha^1, \cdots, x_\alpha^n$ 使得下面四种可能性之一成立:

(i) $U \cap D = \varnothing$;

(ii) $U \subset D$;

(iii) $\alpha(U \cap D) = \alpha(U) \cap \{\boldsymbol{v} = \prec v^1, \cdots, v^n \succ \in \mathbb{R}^n : v^n \geqslant 0\}$;

(iv) $\alpha(U \cap D) = \alpha(U) \cap \{\boldsymbol{v} = \prec v^1, \cdots, v^n \succ \in \mathbb{R}^n : v^k \geqslant 0, \cdots, v^n \geqslant 0\}$.

新颖之处是现在允许了在 $k < n$ 时的可能性 (iv). 当然，只有当 $n > 1$ 时才有这种新的可能性. 假设 $n > 1$ 并看一看 (iv) 会发生什么. 我们将 $\alpha(U \cap \partial D)$ 可以写成位于 $\mathbb{R}^{n-1}$ 中的一些 $(n-1)$ 维子空间中开子集的并，连同一些位于 $n-2$ 维子空间部分的并.

事实上，对 $k \leqslant p \leqslant n$ 设 H_p^k 是由 $v^p = 0$ 给出的 $(n-1)$ 维子空间的一个开子集，定义为

$$H_p^k = \{\boldsymbol{v} : v^k > 0, \cdots, v^p = 0, v^{p+1} > 0, \cdots, v^n > 0\}$$

(见图 10.14). 我们可以写出

$$\alpha(U \cap \partial D) \subset \alpha(U) \cap \{(H_k^k \cup H_{k+1}^k \cup \cdots \cup H_n^k) \cup S\},$$

其中 S 为那些至少有两个 v^p 为零的 ($(n-2)$ 维) 子空间的并.

注意到如果 $x \in U \cap \partial D$ 是使对某个 p 有 $\alpha(x) \in H_p^k$ 的点，则存在一个 x 的型 (iii) 的分图表. 事实上，简单地重排坐标使得 v^p 变成 v^n，即映射 $\mathbb{R}^n \xrightarrow{\varphi} \mathbb{R}^n$ 将 $\prec v^1, \cdots, v^n \succ$ 送到 $\prec w^1, \cdots, w^n \succ$，其中

$$w^i = v^i, i < p, w^i = v^{i+1}, p \leqslant i < n, w^n = v^p.$$

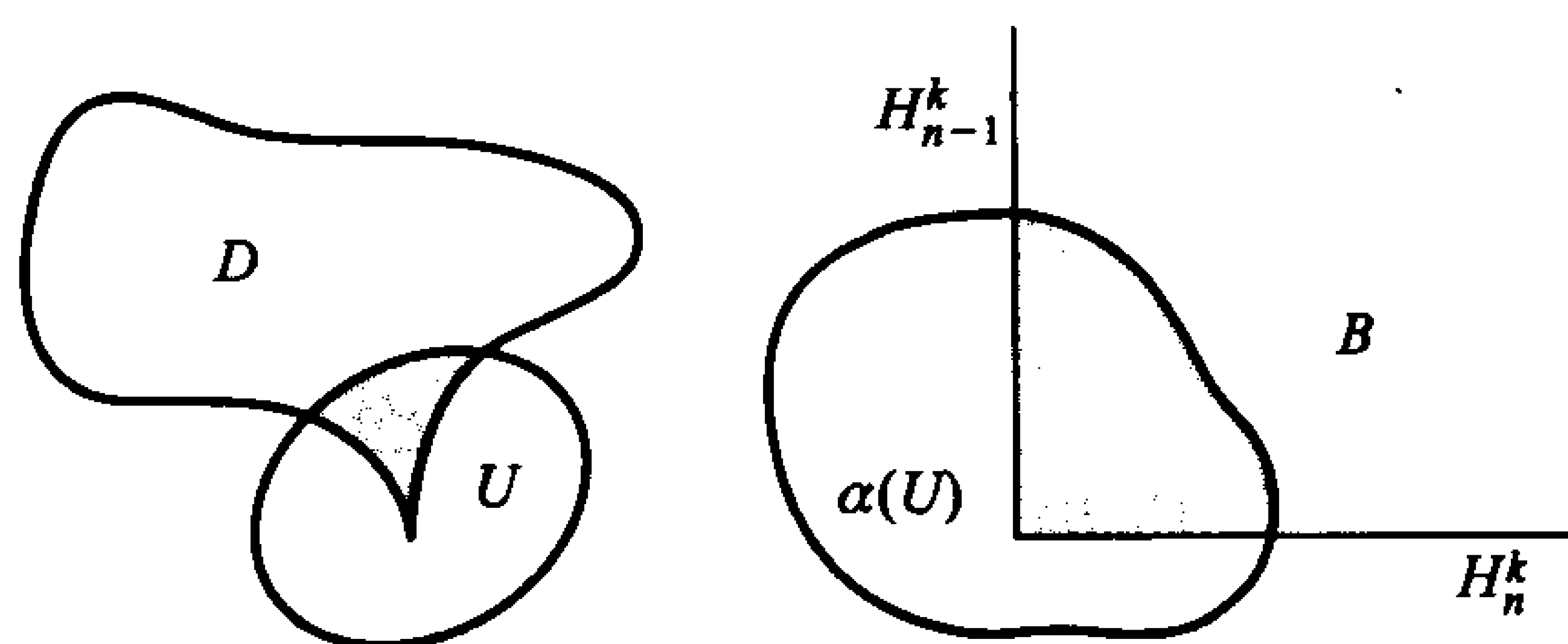


图 10.14

于是, 在 x 的充分小的邻域 U^1, 分图表 $(U^1, \varphi\circ\alpha)$ 具有型 (iii)(见图 10.15).

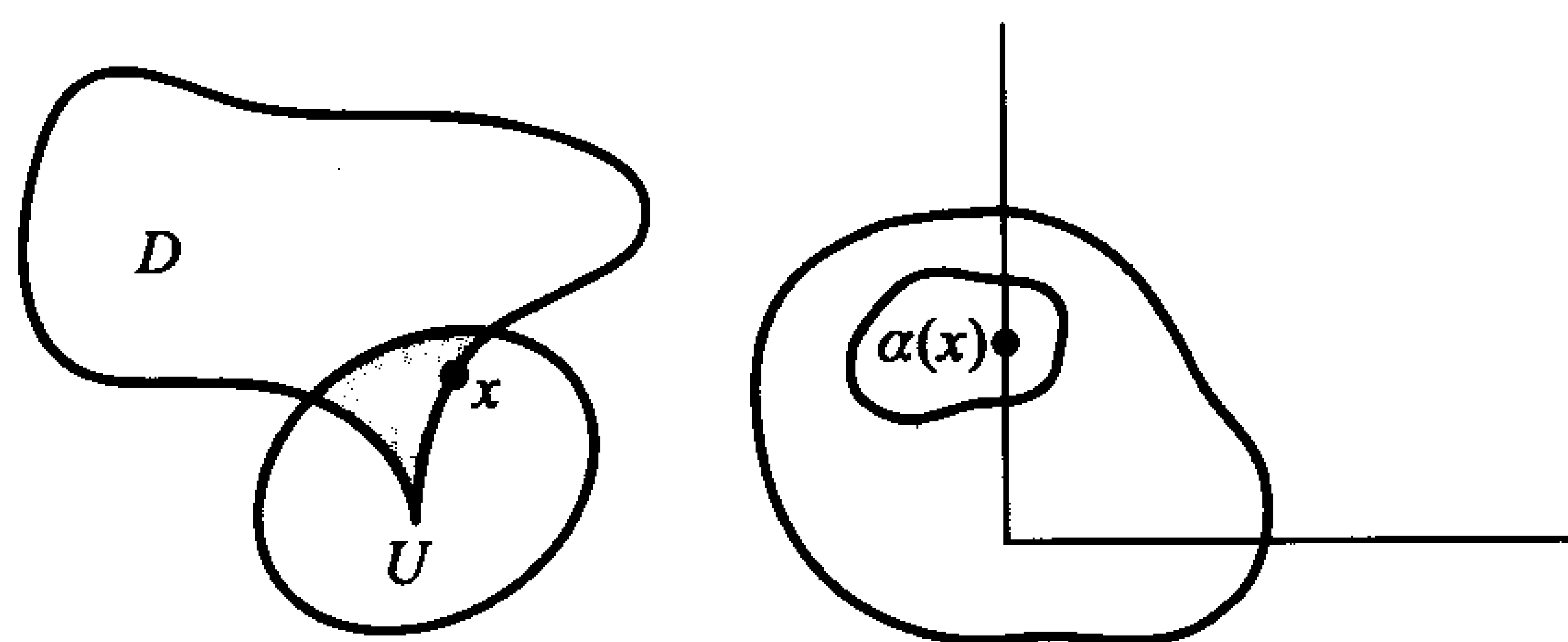


图 10.15

下面我们要看到的是那些 $x\in\partial D$ 的集合构成一个微分流形, 其中每个 x 都有一个型 (iii) 的邻域. 论证完全如前所做. 惟一的差别是这一次的这些点并没有占尽所有 ∂D. 我们以 $\widetilde{\partial D}$ 代表这个集合. 因此 $\widetilde{\partial D}$ 是个流形, 作为集合它不是 ∂D 而只是 ∂D 的"正则"点, 即那些具型 (iii) 的分图表的点.

定理 7.1 (散度定理) 设 M 是个 n 维流形, $D\subset M$ 为一个具几乎正则边界的区域. 设 $\widetilde{\partial D}$ 如上, i 为单射 $\widetilde{\partial D}\to M$. 于是, 对任何 $\rho\in P$ 我们有

$$\int_D \operatorname{div} \prec X,\rho\succ = \int_{\widetilde{\partial D}} \varepsilon_X \rho_X. \tag{7.1}$$

证明 证明如同前面一样的进行. 选一个联结类型 (i) 到 (iv) 的分图表的总图表和一个从属于这个总图表的 1 的分解 $\{g_j\}$. 记 $\rho=\sum g_i\rho$, 而现在有 4 种情形要考虑. 前三种我们已经处理过了.

新的情形是在 ρ 的支集属于 U 而 (U,α) 是型 (iv) 的一个分图表时发生的. 我们必须计算

$$\int_{\alpha(U\cap D)}\sum\frac{\partial X^i\rho_\alpha}{\partial x^i}$$

像前面一样, 和式中那些对应于 $i < k$ 的项对此积分没有作用. 我们现在扩张 $X^i\rho_\alpha$, 使其在 $\alpha(U)$ 之外为零从而定义在整个 $\mathbb{R}^n$, 这也如前面所做. 于是, 对 $k \leqslant i \leqslant n$ 我们有

$$\int_{\alpha(U\cap D)} \frac{\partial X^i\rho_\alpha}{\partial x^i} = \int_B \frac{\partial X^i\rho_\alpha}{\partial x^i},$$

其中 $B = \{v : v^k \geqslant 0, \cdots, v^n \geqslant 0\}$. 写其为累次积分并首先做关于 x^i 的积分, 我们便得到

$$\int_B \frac{\partial X^i\rho_\alpha}{\partial x^i} = \int_{A_i} X^i\rho_\alpha,$$

其中 $A_i \subset \mathbb{R}^{n-1}$ 由

$$A_i = \{\prec v^1, \cdots, v^{i-1}, v^{i+1}, \cdots, v^n \succ : v^k \geqslant 0, \cdots, v^n \geqslant 0\}.$$

注意, A_i 与 H_i^k 相差一个在 $\mathbb{R}^{n-1}$ 中一个容度零的集合 (即在那里至少有一个 $v^l = 0$, 其中 $k = l \leqslant n$). 因此我们可以用 H_i^k 代替 A_i 进行积分. 对 $k \leqslant i \leqslant n$ 求和, 得到

$$\int_{\alpha(D\cap U)} \sum \frac{\partial X^i\rho_\alpha}{\partial x^i} = \sum_{i=k}^{n} \int_{H_i^k} X^i\rho_\alpha,$$

这恰恰是定理 7.1 对情形 (iv) 的断言. □

我们必须指出, 甚至定理 7.1 也没有覆盖那些能用散度定理的所有的情形. 例如, 在平面上定理 7.1 可应用在 D 是个三角形的情形 (见图 10.16). 这是因为我们可以 "撑开" 每个角为直角 (事实上, 可以通过 $\mathbb{R}^2$ 的变量线性变换做到这点.)(见图 10.17).

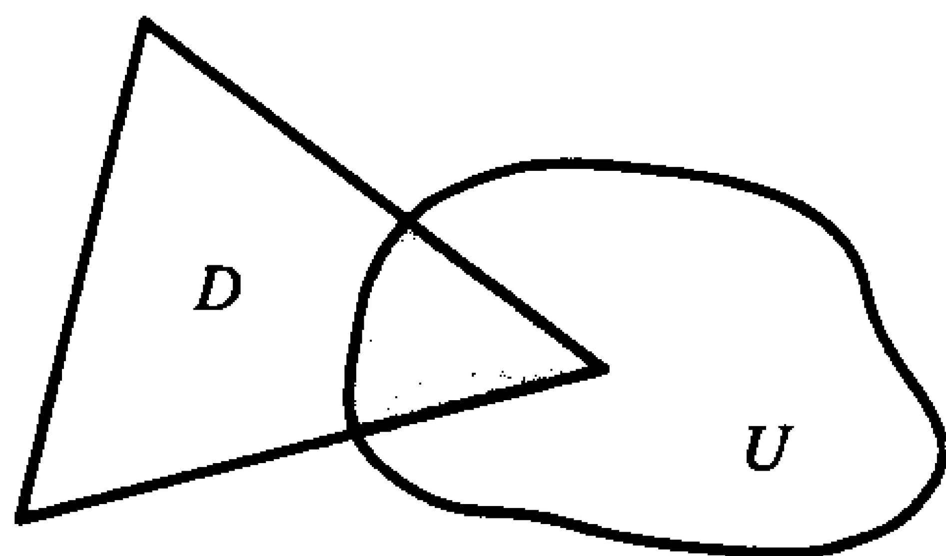


图 10.16

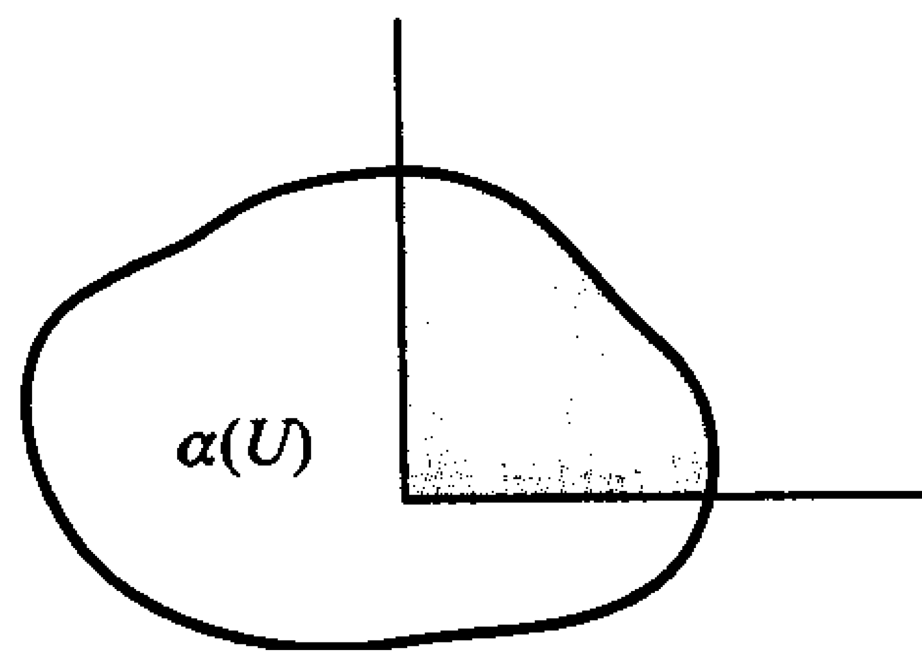


图 10.17

定理 7.1 也不能用在像在图 10.18 中那样一个四边形上, 这是因为没有一个 C^1 变换能把一个大于 π 的角转换为一个小于 π 的角 (由于它的雅可比在拐角处

必须把直线变成直线的缘故). 定理 7.1 不能直接用上去. 但是, 我们可以把四边形写成两个三角形的并, 再应用定理 7.1 到每个三角形上, 并留意到每个三角形来自公共边界的贡献相互抵消了. 因此散度定理的确可用到我们的四边形上.

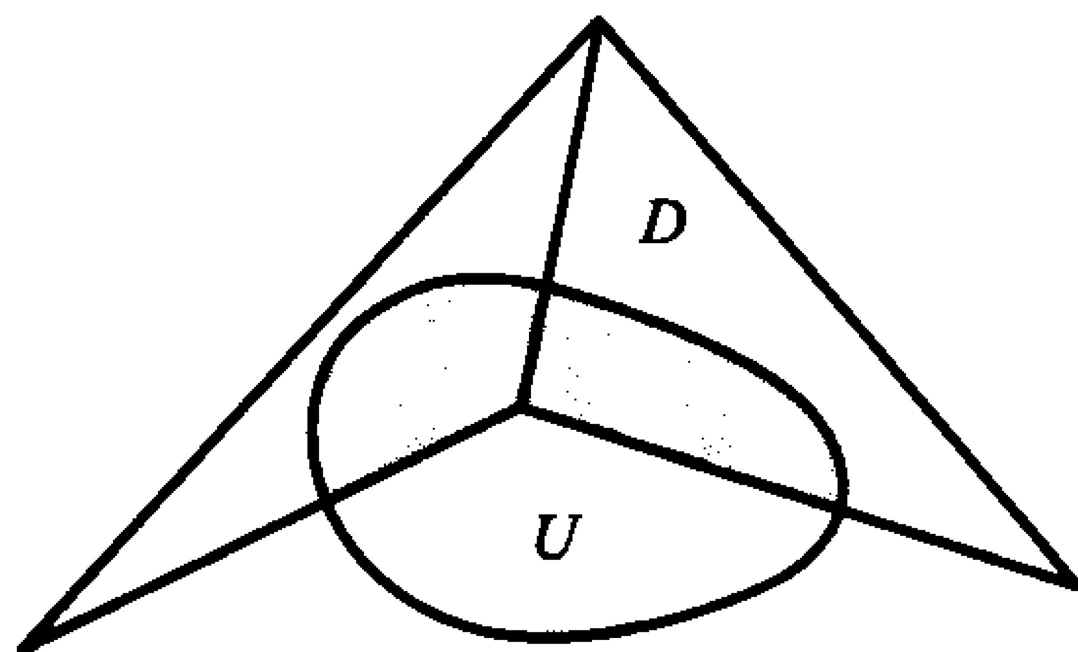


图 10.18

这种步骤在相当广泛的情形都可奏效. 事实上, 在本书需要用到散度定理的地方它都是行得通的, 不管定理 7.1 直接应用还是将它通过对我们区域的有限次细分再随之而来的一种有限度的论证, 就能做到. 但是, 我们将不正式叙述一个包括了所有这些情形的一般性定理; 在每个场合该如何进行是清楚的.

习题

在欧几里得空间, 我们以 $\mathrm{div}X$ 替代 $\mathrm{div} \prec X, \rho \succ$, 这里的 ρ 是欧几里得体积密度.

7.1 设 x, y, z 为 $\mathbb{E}^3$ 的直角坐标. 设向量场 X 由

$$X = r^2\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}\right)$$

给出, 其中 $r^2 = x^2 + y^2 + z^2$. 直接由两边积分证明

$$\int_S (X, \boldsymbol{n})dA = \int_B \mathrm{div}X.$$

这里的 B 是中心在原点而 S 为其边界的球.

7.2 设向量场 Y 由

$$Y = Y_r n_r + Y_\theta n_\theta + Y_\varphi n_\varphi$$

给出, 所用的是 $\mathbb{E}^3$ 中的极 "坐标" r, θ, φ, 其中的 n_r, n_θ, n_φ 是在方向 $\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \varphi}$ 上各自的单位向量.

证明

$$\mathrm{div}Y = \frac{1}{r^2\sin\varphi}\left\{\frac{\partial}{\partial r}(r^2\sin\varphi Y_r) + \frac{\partial}{\partial \theta}(rY_\theta) + \frac{\partial}{\partial \varphi}(r\sin\varphi Y_\varphi)\right\}.$$

7.3 用平面的极坐标计算一个向量场的散度.

7.4 用 $\mathbb{E}^3$ 中的柱坐标计算一个向量场的散度.

7.5 设 σ 为单位球面 S^2 上的体积 (面积) 密度. 用在球面上的坐标 θ, φ (极坐标) 计算 $\mathrm{div}\sigma X$.

第十一章　外微积分

设 M 是个微分流形, ω 为 M 上的一个线性微分形式. 对任何可微曲线 $C:[a,b]\longrightarrow M$ 考虑积分 $\int_a^b\langle C'(t),\omega_{C(t)}\rangle dt$. 设 $[c,d]\longrightarrow[a,b]$ 为由 $s\longmapsto t(s)$ 给出的一个可微映射. 曲线 $B:[c,d]\longrightarrow M$ 由 $B(s)=C(t(s))$ 给出, 它满足

$$B'(s)=t'(s)C'(t(s)).$$

因此如果 $t'(s)>0$ 对所有 s 成立, 则

$$\int_c^d\langle B'(s),\omega_{B(s)}\rangle ds=\int_a^b\langle C'(t),\omega_{C(t)}\rangle dt.$$

从而一个线性微分形式是某种我们可以在 M 的"定向"曲线上积分并与参数化无关的东西. 这一章中我们将引进那些可以在 M 的"定向 k 维曲面"上进行积分的对象并研究它们的性质.

11.1　外微分形式

我们曾定义一个线性微分形式为一个规则, 它对每个 $x\in M$ 指定 $T_x^*(M)$ 中一个元素. 我们将 $T_x^*(M)$ 看作 $\mathfrak{a}^1(T_x(M))$. 鉴于此, 我们对此定义作如下推广. 我们说 M 上一个 q 次外微分形式是指一个规则, 它对每个 $x\in M$ 指定 $\mathfrak{a}^q(T_x(M))$

中一个元素. 如果 ω 是个 q 次外形式, (U,α) 是个分图表, 则由于 α 对 $x\in U$ 将每个 $T_x(M)$ 等同于 V, 我们便得到一个 $\mathfrak{a}^q(V)$ - 值函数 ω_α, 在 $\alpha(U)$ 上由

$$\omega_\alpha(v)(\xi_\alpha^1,\cdots,\xi_\alpha^q)=\omega(x)(\xi^1,\cdots,\xi^q)$$

定义, 其中 $v=\alpha(x),\xi^1,\cdots,\xi^q\in T_x(M)$.

容易写出转移法则. 事实上, 如果 (W,β) 是第二个分图表, 我们则有

$$\omega_\beta(v)(\xi_\beta^1,\cdots,\xi_\beta^q)=\omega(x)(\xi^1,\cdots,\xi^q)=\omega_\alpha(\alpha(x))(\xi_\alpha^1,\cdots,\xi_\alpha^q),$$

或者, 由于 $\xi_\beta=J_{\beta\circ\alpha}+(\alpha(x))(\xi_\alpha),\xi\in T_x(M)$, 我们得到

$$\omega_\alpha(v)(\xi_\alpha^1,\cdots,\xi_\alpha^q)=\omega_\beta(\beta\circ\alpha^{-1}(v))(J_{\beta\circ\alpha^{-1}}(v)\xi_\alpha^1,\cdots,J_{\beta\circ\alpha^{-1}}(v)\xi_\alpha^q). \tag{1.1}$$

为了把 (1.1) 式写成不那么繁琐的形式, 我们引进下面的记号. 设 V_1 和 V_2 为向量空间, $l:V_1\longrightarrow V_2$ 为线性映射. 我们定义 $\mathfrak{a}^p(l)$ 为 $\mathfrak{a}^p(V_2)\longrightarrow\mathfrak{a}^p(V_1)$ 的线性映射, 定义为, 对所有 $w\in\mathfrak{a}^p(V_2)$ 及 $v_1,\cdots,v_p\in V_1$ 有:

$$\mathfrak{a}^p(l)(w)(v_1,\cdots,v_p)=w(l(v_1),\cdots,l(v_p)).$$

注意, 在 $\mathfrak{a}^1(V)$ 与 V^* 的等同下, 映射 $\mathfrak{a}^1(l)$ 与映射 $l^*:V_2^*\longrightarrow V_1^*$ 相同. 还要注意, 如果 $w_1\in\mathfrak{a}^p(V_2),w_2\in\mathfrak{a}^q(V_2)$, 则

$$\mathfrak{a}^p(l)w_1\wedge\mathfrak{a}^q(l)w_2=\mathfrak{a}^{p+q}(l)(w_1\wedge w_2). \tag{1.2}$$

这直接由定义推出. 另外, 如果 $l_1:V_1\longrightarrow V_2,l_2:V_2\longrightarrow V_3$, 则

$$\mathfrak{a}^p(l_2\circ l_1)=\mathfrak{a}^p(l_1)\circ\mathfrak{a}^p(l_2). \tag{1.3}$$

显见, 如果 l 可微地依赖于某个参数, 则 $\mathfrak{a}^p(l)$ 对任意 p 也是如此.

现在可以写 (1.1) 式为

$$\omega_\alpha(v)=\mathfrak{a}^q(J_{\beta\circ\alpha^{-1}}(v))\omega_\beta(\beta\circ\alpha^{-1}(v)). \tag{1.1'}$$

由 (1.1′) 式显而易见它与要求 ω_α 是一个光滑函数是相互一致的. 因此如果对所有分图表 (U,α) 所有函数 ω_α 在 $\alpha(U)$ 上为 C^∞, 我们则说 ω 是个光滑的微分形式. 像通常那样, 这只需对一个总图表的所有分图表去验证即可. 我们设 $\wedge^q(M)$ 表示所有 q 次光滑外微分形式的空间.

设 $\omega_1 \in \wedge^p(M), \omega_2 \in \wedge^q(M)$. 定义外微分 $(p+q)$ - 形式 $\omega_1 \wedge \omega_2$ 为, 对所有 $x \in M$ 有

$$(\omega_1 \wedge \omega_2)(x) = \omega_1(x) \wedge \omega_2(x).$$

容易验证 $\omega_1 \wedge \omega_2$ 是个光滑的 $(p+q)$ - 形式. 我们因而得到了在外微分形式上的乘法. 为了使这种公式化陈述完备, 把 M 上的可微函数的空间表示为 $\wedge^0(M)$ 有其方便之处, 另外表示函数 f 与 p - 形式 ω 的乘积为 $f\omega$ 或 $f \wedge \omega$. 这个乘积由对所有 $x \in M$,

$$(f \wedge \omega)(x) = (f\omega)(x) = f(x)\omega(x)$$

给出.

于是我们已经对所有 $0 \leqslant p \leqslant n$ 和 $0 \leqslant q \leqslant n$ 定义了一个乘法, 把 $\omega_1 \in \wedge^p(M)$ 和 $\omega_2 \in \wedge^q(M)$ 送到 $\omega_1 \wedge \omega_2 \in \wedge^{p+q}(M)$ (如果 $p+q > n = \dim M$ 则 $\omega_1 \wedge \omega_2 \equiv 0$). 对于 $\wedge$ - 乘法在反对称张量上的规则也转移到这里, 因此举例说, 有

$$\begin{aligned}
&\text{如果 } \omega_1 \in \wedge^p(M), \omega_2 \in \wedge^q(M), \text{则} \\
&\omega_1 \wedge \omega_2 = (-1)^{pq} \omega_2 \wedge \omega_1, \\
&\omega_1 \wedge (\omega_2 \wedge \omega_3) = (\omega_1 \wedge \omega_2) \wedge \omega_3, \\
&\omega_1 \wedge (\omega_2 + \omega_3) = \omega_1 \wedge \omega_2 + \omega_1 \wedge \omega_3,
\end{aligned}$$

等等.

设 M_1 和 M_2 为微分流形, $\varphi: M_1 \longrightarrow M_2$ 为可微映射. 对每个 $\omega \in \wedge^q(M_2)$ 我们定义形式 $\varphi^*\omega \in \wedge^q(M_1)$ 为

$$\varphi^*\omega(x) = \mathfrak{a}^q(\varphi_{*x})(\omega(\varphi(x))). \tag{1.4}$$

容易验证 $\varphi^*\omega$ 的确是 $\wedge^q(M_1)$ 中一个元素, 即它是一个光滑 q - 形式. 还要注意到, 第九章的 (7.5) 式是 (1.4) 式的一个特例, 即 $q=1$ 的情形.(如果我们约定 $\mathfrak{a}^0(l) = \mathrm{id}$, 则 (1.4) 式的 $q=0$ 便是函数拉回的规则.)

由 (1.4) 式推出 φ^* 是线性的, 即

$$\varphi^*(\omega_1 + \omega_2) = \varphi^*(\omega_1) + \varphi^*(\omega_2), \tag{1.5}$$

又从 (1.2) 式知

$$\varphi^*(\omega_1 \wedge \omega_2) = \varphi^*(\omega_1) \wedge \varphi^*(\omega_2). \tag{1.6}$$

如果 φ 是流形 M 上的一个单参数群, 其无穷小生成元为 X, 则我们可以证明对任意 $\omega \in \wedge^q(M)$,

$$\lim_{t\to 0}\frac{\varphi_t^*\omega-\omega}{t}=D_X\omega$$

存在. 这个极限的存在性的证明是直截了当的, 我们略去它. 在第 3 节中我们要推导出一个可以简单计算 $D_X\omega$ 的有用公式.

我们现在来看看如何用局部坐标计算 $\wedge^q(M)$. 设 (U,α) 为 M 的一个分图表, 坐标为 $x^1,\cdots,x^n$. 于是 $dx^i \in \wedge^1(U)$(在这里我们用 $\wedge^q(U)$ 表示定义在 U 上的可微 q - 形式的集合). 对任意 $i_1,\cdots,i_q$, 形式 $dx^{i_1}\wedge\cdots\wedge dx^{i_q}$ 属于 $\wedge^q(U)$, 并对每个 $x \in U$, 形式

$$\{(dx^{i_1}\wedge\cdots\wedge dx^{i_q})(x)\}_{i_1<\cdots<i_q}$$

构成 $\mathfrak{a}^q(T_x(M))$ 的一组基. 由此推出, 每个 q 次的外微分形式 ω, 如果定义在 U 上则可写为

$$\omega=\sum_{i_1<\cdots<i_q} a_{i_1\cdots i_q}dx^{i_1}\wedge\cdots\wedge dx^{i_q}, \tag{1.7}$$

其中的这些 a 是函数, 就是说, 对所有 $x \in U$ 有

$$\omega(x)=\sum_{i_1<\cdots<i_q} a_{i_1,\cdots,i_q}(x)(dx^{i_1}\wedge\cdots\wedge dx^{i_q})(x).$$

容易看出 $\omega \in \wedge^q(U)$ 的充要条件是在 U 上的所有函数 $a_{i_1,\cdots,i_q}$ 是 C^∞ 函数.

如果 (W,β) 是第二个分图表, 坐标为 $y^1,\cdots,y^n$, 而

$$\omega=\sum b_{j_1,\cdots,j_q}dy^{j_1}\wedge\cdots\wedge dy^{j_q}, \tag{1.8}$$

则容易算出在 $U\cap W$ 上联系这些 b 和 a 的转移法则.

事实上, 在 $U\cap W$ 上我们有

$$dy^j=\sum\frac{\partial y^j}{\partial x^i}dx^i, \tag{1.9}$$

其中 $y^j=y^j(x^1,\cdots,x^n)$. 于是我们全部要做的事就是把 (1.9) 式代入 (1.8) 式并合并 $dx^{i_1}\wedge\cdots\wedge dx^{i_q}$ 的系数. 例如, 如果 $q=2$, 我们则有

$$\begin{aligned}\omega&=\sum_{j_1<j_2} b_{j_1j_2}dy^{j_1}\wedge dy^{j_2}\\&=\sum_{j_1<j_2} b_{j_1j_2}\left(\frac{\partial y^{j_1}}{\partial x^1}dx^1+\cdots+\frac{\partial y^{j_1}}{\partial x^n}dx^n\right)\wedge\left(\frac{\partial y^{j_2}}{\partial x^1}dx^1+\cdots+\frac{\partial y^{j_2}}{\partial x^n}dx^n\right).\end{aligned}$$

如果合并 $dx^{i_1}\wedge dx^{i_2}$ 的系数 (记住 $\wedge$ - 乘法是反交换的), 我们则有

$$\omega=\sum_{i_1<i_2}\left[\sum_{j_1<j_2}b_{j_1j_2}\left(\frac{\partial y^{j_1}}{\partial x^{i_1}}\frac{\partial y^{j_2}}{\partial x^{i_2}}-\frac{\partial y^{j_2}}{\partial x^{i_1}}\frac{\partial y^{j_1}}{\partial x^{i_2}}\right)\right]dx^{i_1}\wedge dx^{i_2}.$$

于是

$$a_{i_1i_2}=\sum_{j_1<j_2}b_{j_1j_2}\det\begin{bmatrix}\dfrac{\partial y^{j_1}}{\partial x^{i_1}} & \dfrac{\partial y^{j_2}}{\partial x^{i_1}}\\ \dfrac{\partial y^{j_1}}{\partial x^{i_2}} & \dfrac{\partial y^{j_2}}{\partial x^{i_2}}\end{bmatrix}. \tag{1.10}$$

虽然 (1.10) 式看起来有点令人生畏, 但关键之处在于人们必须完全记住的是 (1.9) 式和 $\wedge$ - 乘法的法则. 对于一般的 q 同样的论证给出

$$a_{i_1,\cdots,i_q}=\sum_{j_1<\cdots<j_q}b_{j_1,\cdots,j_q}\det\begin{bmatrix}\dfrac{\partial y^{j_1}}{\partial x^{i_1}} & \cdots & \dfrac{\partial y^{j_q}}{\partial x^{i_1}}\\ \vdots & & \vdots\\ \dfrac{\partial y^{j_1}}{\partial x^{i_q}} & \cdots & \dfrac{\partial y^{j_q}}{\partial x^{i_q}}\end{bmatrix}. \tag{1.11}$$

拉回的公式有完全一样的样子. 设 φ: $M_1\longrightarrow M_2$ 为可微映射, 并假定 (U,α) 和 (W,β) 为相容的分图表, 其中 $x^1,\cdots,x^m$ 为 (U,α) 的坐标, $y^1,\cdots,y^n$ 为 (W,β) 的坐标. 于是我们知道 $y^i\circ\varphi$ 是 U 上的函数, 因而可以写为

$$y^j\circ\varphi=y^j(x^1,\cdots,x^m).$$

由于 $\varphi^*dy^j=d(y^j\circ\varphi)$, 我们有

$$\varphi^*(dy^j)=\sum\frac{\partial y^j}{\partial x^i}dx^i. \tag{1.12}$$

如果

$$\omega=\sum_{j_1<\cdots<j_q}b_{j_1,\cdots,j_q}dy^{j_1}\wedge\cdots\wedge dy^{j_q}\in\wedge^q(W),$$

则由 (1.5) 和 (1.6) 式有

$$\varphi^*\omega=\sum_{j_1<\cdots<j_q}(b_{j_1,\cdots,j_q}\circ\varphi)(\varphi^*dy^{j_1}\wedge\cdots\wedge\varphi^*dy^{j_q}). \tag{1.13}$$

(1.13) 式用这些 dx 表出的表达式可由将 (1.12) 式代入 (1.13) 式进行计算并合并系数项. 自然, 其答案看起来就是前面做过的那个样子. 如果

$$\varphi^*(\omega)=\sum_{i_1<\cdots<i_q}a_{i_1,\cdots,i_q}dx^{i_1}\wedge\cdots\wedge dx^{i_q},$$

则这些 a 由

$$a_{i_1,\cdots,i_q}=\sum_{j_1<\cdots<j_q}(b_{j_1,\cdots,j_q}\circ\varphi)\det\begin{bmatrix}\dfrac{\partial y^{j_1}}{\partial x^{i_1}} & \cdots & \dfrac{\partial y^{j_q}}{\partial x^{i_1}}\\ \vdots & & \vdots\\ \dfrac{\partial y^{j_1}}{\partial x^{i_q}} & \cdots & \dfrac{\partial y^{j_q}}{\partial x^{i_q}}\end{bmatrix}. \tag{1.14}$$

我们再次强调, 没有必要去记住一个像 (1.14) 式这种样子复杂的公式; 用等式 (1.5),(1.6) 和 (1.12)(当然还有 $\wedge$ - 乘法规则) 就够了. 在许多情形直接进行替换比起用 (1.14) 式要方便很多.

11.2 定向流形和外微分形式的积分

设 M 为 n 维流形, (U,α) 和 (W,β) 为 M 上两个分别具有坐标 $x^1,\cdots,x^n$ 和 $y^1,\cdots,y^n$ 的分图表. 设 ω 为一个 n 次的外微分形式. 于是我们可以写为

$$\text{在 } U \text{ 上}:\ \omega=adx^1\wedge\cdots\wedge dx^n,$$
$$\text{在 } W \text{ 上}:\ \omega=bdy^1\wedge\cdots\wedge dy^n,$$

其中函数 a 和 b 在 $U\cap W$ 上由 (1.11) 式相关联, 在目前的情形 $(q=n)$ 成为

$$a=b\det\begin{bmatrix}\dfrac{\partial y^1}{\partial x^1} & \cdots & \dfrac{\partial y^n}{\partial x^1}\\ \vdots & & \vdots\\ \dfrac{\partial y^1}{\partial x^n} & \cdots & \dfrac{\partial y^n}{\partial x^n}\end{bmatrix}$$

或者

$$a_\alpha(\alpha(x))=b_\beta(\beta\circ\alpha^{-1}(\alpha(x)))\det J_{\beta\circ\alpha^{-1}}(\alpha(x)),$$

最后或者写为, 对 $v\in\alpha(U\cap W)$ 有

$$a_\alpha(v)=b_\beta(\beta\circ\alpha^{-1}(v))\det J_{\beta\circ\alpha^{-1}}(v). \tag{2.1}$$

如果 ρ 是 M 上一个密度, 则对于 ρ_α 的转移法则由

$$\rho_\alpha(v)=\rho_\beta(\beta\circ\alpha^{-1}(v))\mid\det J_{\beta\circ\alpha^{-1}}(v)\mid \tag{2.2}$$

给出. 注意, (2.2) 和 (2.1) 式看起来几乎一样; 差别在于在 (2.2) 式中出现了绝对值符号而在 (2.1) 式中没有. 特别, 如果 (U,α) 和 (W,β) 使得 $\det J_{\beta\circ\alpha^{-1}} > 0$, 则 (2.2) 和 (2.1) 式在这一对分图表上便是一致的.

这个评注把我们带到了下面的定义: M 的一个总图表 $\mathfrak{a}$ 被称为定向的是说对 $\mathfrak{a}$ 的任一对分图表 (U,α) 和 (W,β) 我们对所有 $x \in U \cap W$ 有

$$\det J_{\beta\circ\alpha^{-1}}(\alpha(x)) > 0.$$

不能保证在一个给定的流形 M 上存在有一个定向总图表. 事实上不难证明在某些流形上的确不存在定向的总图表.(不具定向总图表的流形的例子是麦比乌斯带.)

我们说一个流形 M 是可定向的是指它有一个定向的总图表.

设 M 是一个可定向流形, 设 $\mathfrak{a}_1$ 和 $\mathfrak{a}_2$ 是两个定向总图表. 称 $\mathfrak{a}_1$ 和 $\mathfrak{a}_2$ 有相同的定向并记为 $\mathfrak{a}_1 \tilde{\circ} \mathfrak{a}_2$ 是说 $\mathfrak{a}_1 \cup \mathfrak{a}_2$ 还是一个定向总图表. 说 $\mathfrak{a}_1 \tilde{\circ} \mathfrak{a}_2$ 意味着对任何 $(U,\alpha) \in \mathfrak{a}_1$ 和任何 $(W,\beta) \in \mathfrak{a}_2$, 我们在 $\alpha(U \cap W)$ 上有

$$\det J_{\beta\circ\alpha^{-1}}(v) > 0.$$

显然 $\tilde{\circ}$ 是个等价关系. 定向总图表的一个等价类被称为 M 的一个定向. 一个可定向流形连同选取的一个定向被称为一个定向流形. 我们以 $\boldsymbol{M}$ 表示一个定向流形. 即 $\boldsymbol{M}$ 是个流形 M 连同一个所选择的定向. 因此一个定向一维流形在每个点有一选择的方向 (图 11.1); 一个定向的二维流形有一个顺时针还是反时针方向的概念 (图 11.2); 在一个定向三维流形上的任一点, 我们可以区分是右手系还是左手系.

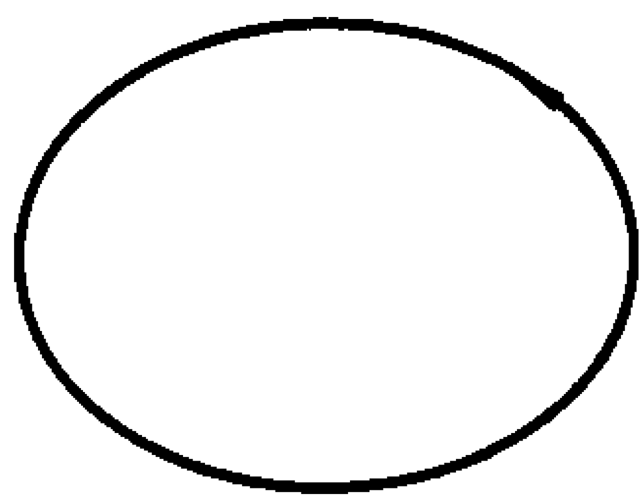


图 11.1

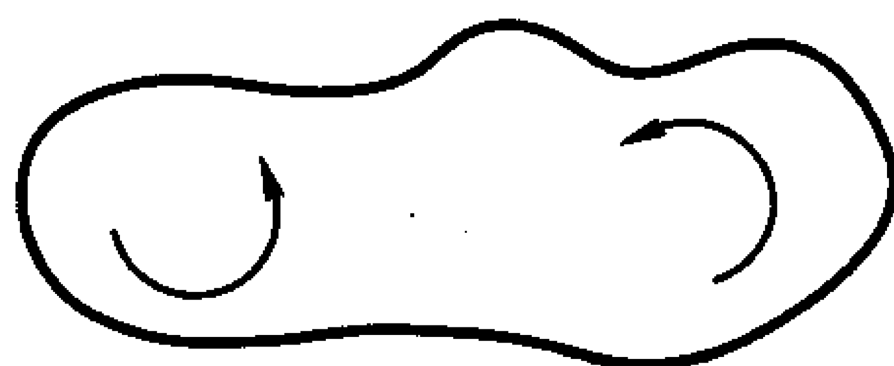


图 11.2

一般地, 设 $\boldsymbol{M}$ 是个定向流形, (U,α) 是 M 的一个分图表, 坐标为 $x^1,\cdots,x^n$. 我们说 (U,α) 是一个正分图表是指对定义 (或属于) 这个定向的任一总图表中任何一个分图表 (W,β) 有 $\det J_{\beta\circ\alpha^{-1}} > 0$.(自然, 只要对属于一个固定的定义定向

的总图表去检验所有的 (W,β) 就可以了.) 注意, 如果 U 是连通的, 那么, 如果 (U,α) 不是正的, 则分图表 (U,α^1) 便是正的, 其中如果 $\alpha(x) = \prec v^1, v^2, \cdots, v^n \succ$,

$$\alpha^1(x) = \prec -v^1, v^2, \cdots, v^n \succ .$$

如果对 $(W,\beta),(U,\alpha)$ 有 $\det J_{\beta\circ\alpha^{-1}} < 0$, 我们则说 (U,α) 是个负分图表, 这里的 (W,β) 是属于定义此定向的一个总图表的所有分图表. (从而如果 U 连通, 则 (U,α) 必定或正或负.)

我们现在回到比较 (2.1) 与 (2.2) 式的最初的考察上.

命题 2.1 设 M 为定向 n 维流形. 我们可以把 n 次外微分形式等同于密度, 这个等同把形式 ω 送到密度 ρ^ω, 其中对任一坐标为 $x^1,\cdots,x^n$ 的正分图表 (U,α), 函数 ρ^ω_α 由在 U 上

$$\omega = \rho^\omega_\alpha(\alpha(\cdot))dx^1 \wedge \cdots \wedge dx^n \tag{2.3}$$

决定.

(2.3) 式的另一种写法是

$$\omega\left(\frac{\partial}{\partial x^1},\cdots,\frac{\partial}{\partial x^n}\right) = \rho\left(\frac{\partial}{\partial x^1},\cdots,\frac{\partial}{\partial x^n}\right). \tag{2.3'}$$

换句话说, 如果在 U 上, $\omega = a dx^1 \wedge \cdots \wedge dx^n$, 则 $\rho^\omega_\alpha(v) = a_\alpha$. 因为对所有属于一个正总图表的一对分图表, (2.2) 式约化成为 (2.1) 式, 故 ρ^ω 确实是一个密度.

这个等同显然是可加的, 即

$$\rho^{\omega_1+\omega_2} = \rho^{\omega_1} + \rho^{\omega_2}, \tag{2.4}$$

而且对任意函数 f, 有

$$\rho^{f\omega} = f\rho^\omega. \tag{2.5}$$

另外, 如果 $\omega(x) = 0$ 则在 $x, \rho^\omega = 0$. 像通常那样, 一个微分形式的支集表示是使 $\omega(x) \neq 0$ 的那些 x 的集合的闭包. 我们说一个 n - 形式 ω 为局部绝对可积是指密度 ρ^ω 是局部绝对可积的. 注意, 说 ω 是局部绝对可积意味着对任何具有坐标 $x^1,\cdots,x^n$ 的某个 $\mathfrak{a}$ 中的分图表 (U,α) 如果在 U 上

$$\omega = a dx^1 \wedge \cdots \wedge dx^n,$$

则函数 $a_\alpha = a\circ\alpha^{-1}$ 是在 $\alpha(U)$ 上的绝对可积函数. 设 $\Gamma(M)$ 表示具紧支集的绝对可积的 n - 形式的空间, 显然 $\Gamma(M)$ 是个向量空间且对 f 为 (有界) 可度函数, $\omega \in \Gamma(M)$ 时有 $f\omega \in \Gamma(M)$. 作为命题 2.1 和第十章定理 3.1 的推论, 我们可以叙述下面的

定理 2.1 设 M 为定向流形. 于是在 $\Gamma(M)$ 上存在一个惟一的线性函数 $\int$, 满足下面的条件: 如果 $\operatorname{supp}\omega \subset U$, 则

$$\int \omega = \int_{\alpha(U)} a_\alpha, \tag{2.6}$$

其中 (U,α) 是个有坐标 $x^1,\cdots,x^n$ 的一个正分图表, $\omega = a dx^1\wedge\cdots\wedge dx^n$.

注意, 对所有 $\omega\in\Gamma(M)$ 我们可以写出

$$\int \omega = \int \rho^\omega. \tag{2.7}$$

计算 $\int\omega$ 的方法现在非常简单. 我们把 ω 分成小段, 使得每段位于某个 U 内. (在此过程中可以忽略密度零的集合不计.) 如果 $\operatorname{supp}\omega\subset U$, 并且 (U,α) 为正分图表, 我们则可表示 ω 为

$$\omega = a dx^1\wedge\cdots\wedge dx^n.$$

如果 a 由 $a = a_\alpha(x^1,\cdots,x^n)$ 给出, 我们则可在 $\mathbb{R}^n$ 上对函数 a_α 进行积分. 计算是机械性的. 但有一点是必须验证的, 即分图表 (U,α) 为正. 如果为负则 $\int\omega$ 由 $-\int a_\alpha$ 给出.

设 M_1 为定向 q 维流形, $\varphi\colon M_1 \longrightarrow M_2$ 为可微映射, $\omega\in\wedge^q(M_2)$. 于是对任意可度紧集 $A\subset M_1$, 形式 $e_A\varphi^*(\omega)$ 属于 $\Gamma(M_1)$, 故我们可以考虑它的积分. 这个积分有时用 $\int_{\varphi(A)}\omega$ 表示; 就是说, 定义

$$\int_{\varphi(A)}\omega = \int_{M_1} e_A\varphi^*\omega. \tag{2.8}$$

如果我们把 $\varphi(A)$ 看作 M_2 中一个“定向的 q 维曲面”, 我们则看到 $\wedge^q(M_2)$ 的这些元素是可以在这样的“曲面”上进行积分的对象.(在 $q=1$ 时自然称其为“曲线”.)

让我们用一些例子来作解释. 假设 $M_2=\mathbb{R}^n$, 设 $A\subset\mathbb{R}^1$ 为区间 $a\leqslant t\leqslant b$. 设 $x^1,\cdots,x^n$ 为 $\mathbb{R}^n$ 的坐标, 并设 $\omega = a^1dx^1+\cdots+a^ndx^n$. 我们将 $\mathbb{R}^1$ 看作是一

个定向流形, 它上面的恒等分图表是正的 (其坐标为 t). 如果 $C:\mathbb{R}^1\longrightarrow\mathbb{R}^n$ 是可微曲线 (图 11.3), 则

$$\begin{aligned}\int_{C([a,b])}\omega=\int e_{[a,b]}C^*\omega&=\int_a^b\left(a^1\frac{dx^1}{dt}+\cdots+a^n\frac{dx^n}{dt}\right)dt\\&=\int_a^b\prec C't,\omega\succ dt.\end{aligned}\tag{2.9}$$

从这个最后面的表达式我们看到要使 $\int_{C([a,b])}\omega$ 有意义, C 不必是处处可微的. 事实上, 如果除去有限个点外 C 在 $\mathbb{R}$ 上是处处可微的, 并且如果 $C'(t)$ 总是有界的 (在看成 $\mathbb{R}^n$ 中的元素时), 则函数 $\prec C'(\cdot),\omega\succ$ 除去一个密度为零的集合外处处有定义并且有界. 因此 $C^*(\omega)$ 是一个可度的密度, 从而 (2.9) 式依然有意义. 现在此曲线有了拐角 (见图 11.4).

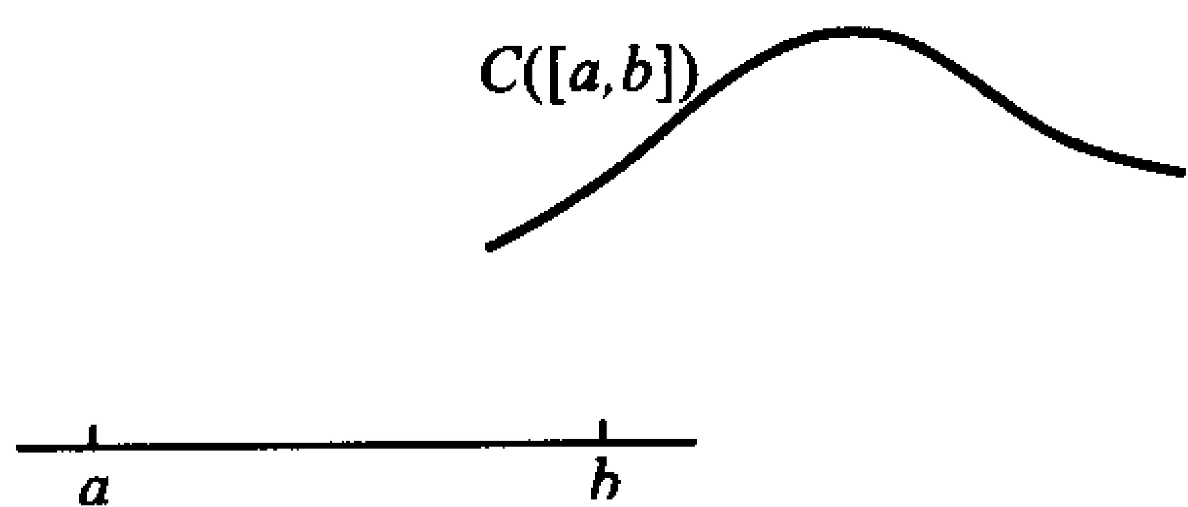


图 11.3

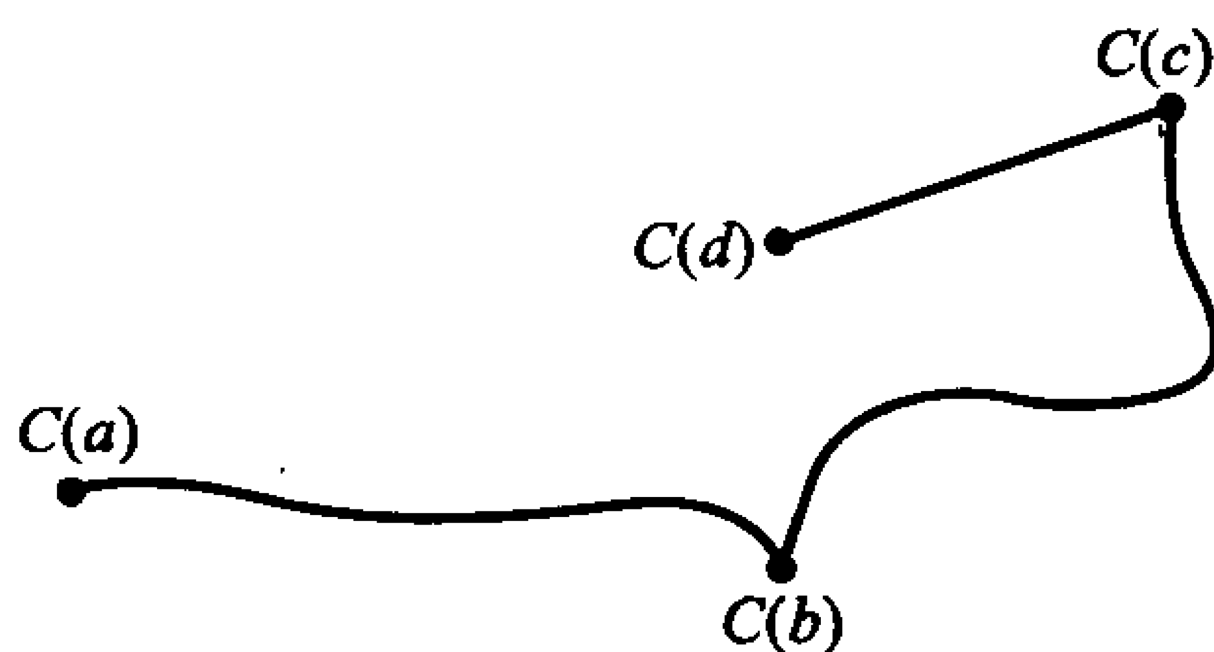


图 11.4

应当看到, 如果 $\omega=df$(并且如果 C 为连续), 则

$$\int_{C([a,b])}df=\int_{C([a,b])}d(f\circ C)=\int_a^b(f\circ C)'=f(C(b))-f(C(a)).\tag{2.10}$$

这时, 此积分不依赖于特定的曲线 C 而依赖于它的端点. 一般情况下, $\int_C\omega$ 依赖于曲线 C. 在第 5 节我们将得到它与 C 无关的条件.

在下一个例题中, 设 $M_2=\mathbb{R}^3, M_1=U\subset\mathbb{R}^2, (u,v)$ 为 $\mathbb{R}^2$ 的欧几里得坐标, x,y,z 为 $\mathbb{R}^3$ 的欧几里得坐标. 设

$$\omega=Pdx\wedge dy+Qdx\wedge dz+Rdy\wedge dz$$

为 $\wedge^2(\mathbb{R}^3)$ 中的元素. 如果 $\varphi: U\longrightarrow\mathbb{R}^3$ 由函数 $x(u,v),y(u,v)$ 和 $z(u,v)$ 给出, 则

对 $A\subset U$ 有

$$\begin{aligned}\int_{\varphi(A)}\omega &= \int e_A\varphi^*\omega = \int e_A\varphi^*(Pdx\wedge dy+Qdx\wedge dz+Rdy\wedge dz)\\ &= \int_A\left[(P\circ\varphi)\left(\frac{\partial x}{\partial u}\frac{\partial y}{\partial v}-\frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right)+(Q\circ\varphi)\left(\frac{\partial x}{\partial u}\frac{\partial z}{\partial v}-\frac{\partial z}{\partial u}\frac{\partial x}{\partial v}\right)\right.\\ &\quad+\left.(R\circ\varphi)\left(\frac{\partial y}{\partial u}\frac{\partial z}{\partial v}-\frac{\partial z}{\partial u}\frac{\partial y}{\partial v}\right)\right].\end{aligned}$$

我们以对黎曼度量的体积密度做另一次观察来结束本节, 这一次的对象是定向流形. 如果 M 是有一个黎曼度量的定向流形, 于是体积密度 σ 对应于一个 n - 形式 Ω. 由我们对这个对应的规则, 如果 (U,α) 是个坐标为 $x^1,\cdots,x^n$ 的正分图表, 则

$$\Omega = adx^1\wedge\cdots\wedge dx^n,$$

又由第十章的 (4.1) 式, 其中的 $a(x)=|\det(g_{ij})|^{\frac{1}{2}}$ 是在 $T_x(M)$ 中由

$$\frac{\partial}{\partial x^1}(x),\cdots,\frac{\partial}{\partial x^n}(x)$$

张成的平行六面体的体积. 设 $e_1(x),\cdots,e_n(x)$ 为 $T_x(M)$ 的法正交基 (相对于此黎曼度量给出的内积). 于是

$$|\det(g_{ij})|^{\frac{1}{2}} = \left|\det\left[\left(\frac{\partial}{\partial x_i},e_j\right)\right]\right| = |\det A|,$$

其中 $A=\left(\frac{\partial}{\partial x^i},e_j\right)$ 是将 $e_j\longmapsto\frac{\partial}{\partial x^j}$ 的线性变换的矩阵. 如果 $\omega^1(x),\cdots,\omega^n(x)$ 为 $\{e_j\}$ 的对偶基, 则

$$\omega^1(x)\wedge\cdots\wedge\omega^n(x) = \det A dx^1(x)\wedge\cdots\wedge dx^n(x).$$

现 $\omega^1(x),\cdots,\omega^n(x)$ 可以是 $T_x^*(M)$ 中任一组法正交基. [$T_x^*(M)$ 由于是内积空间 $T_x(M)$ 的对偶空间, 故也有一个内积.] 我们于是有如下结果: 如果 $\omega^1,\cdots,\omega^n$ 为线性微分形式使得对每个 $x\in M,\omega^1(x),\cdots,\omega^n(x)$ 为 $T_x^*(M)$ 的一组法正交基, 则

$$\Omega = \pm\omega^1\wedge\cdots\wedge\omega^n.$$

如果我们知道 $\omega^1\wedge\cdots\wedge\omega^n$ 是 $dx^1\wedge\cdots\wedge dx^n$ 的正倍数, 我们可以写成

$$\Omega = \omega^1\wedge\cdots\wedge\omega^n. \tag{2.11}$$

我们总能够在 U 上找到这样的 $\omega^1,\cdots,\omega^n$ 吗？回答为"是"：我们可以对 $dx^1,\cdots,dx^n$ 应用法正交化程序. 就是说, 我们令

$$\omega^1=\frac{dx^1}{\|dx^1\|}，\text{其中}\ \|dx^1\|(x)=\|dx^1(x)\|>0\ \text{是}\ U\ \text{上的}\ C^\infty\ \text{函数},$$

$$\omega^2=\frac{dx^2-(dx^2,\omega^1)\omega^1}{\|dx^2-(dx^2,\omega^1)\omega^1\|},$$

$$\vdots$$

联系这些 dx 和这些 ω 的矩阵由 C^∞ 函数组成, 故 $\omega^i\in\wedge^1(U)$. 进一步看, 它是在对角线上为正元的一个三角形矩阵, 故其行列式为正. 因而我们已构造出了所要的形式 $\omega^1,\cdots,\omega^n$, 从而 (2.11) 式成立. 举例说, 由第九章的等式 (9.10) 得到 $d\theta,\sin\theta d\varphi$ 构成 $T_x(S^2)$ 的法正交基, 它对所有 $x\in S^2$ 成立 (除去北极和南极). 如果我们在 S^2 上选取定向, 使得 θ,φ 构成一个正的分图表, 则体积形式由

$$\Omega=\sin\theta d\theta\wedge d\varphi$$

给出.

11.3 算子 d

对每个函数 f 我们有一个关联的线性微分形式 df. 我们因此能将 d 看作从 $\wedge^0(M)$ 到 $\wedge^1(M)$ 的映射. 就其本身而论, 它是线性的, 并满足

$$d(f_1f_2)=f_2df_1+f_1df_2.$$

我们现在试图对 $k>0$ 同样定义一个 $d:\wedge^k(M)\longrightarrow\wedge^{k+1}(M)$. 我们要求 d 是线性的并关于乘法满足某种能推广对 $d(f_1f_2)$ 的等式. 我们加上的这个条件是

$$d(\omega_1\wedge\omega_2)=d\omega_1\wedge\omega_2+(-1)^p\omega_1\wedge d\omega_2,$$

其中 ω_1 是个 p 次形式. 因子 $(-1)^p$ 是出于 $\wedge$ 的反交换的性质. 读者应该检验 d 与这个规律的一致性, 至少在 ω_1 为 p 次, ω_2 为 q 次时验证 $d(\omega_1\wedge\omega_2)=(-1)^{pq}d(\omega_2\wedge\omega_1)$.

我们要再加上对 d 的一个条件以使 d 能被惟一地确定. 对这个条件 (它处在关于 d 的定义的核心地位) 需要介绍一些想法. 设 f 为可微函数, $C:I\longrightarrow M$ 为可微曲线. 对任意点 $a,b\in I$, 微积分基本定理表明

$$f(C(b)) - f(C(a)) = \int_a^b \frac{d(f \circ C)}{dt} dt = \int_a^b C^* df.$$

我们可以把 b 和 a(及其所带的 ± 号) 看作是区间 $[a, b]$ 的“定向边界”. 我们约定“积分” $\wedge^0(p)$ 中一个元素就是此函数在点 p 的取值. 这样一来, 上面的等式所说的便是 f 的“拉回”在“边界”上的积分, 即 $f(b) - f(a)$, 等于 df 的“拉回”在 $[a, b]$ 上的积分. 从某种意义上, 我们想要能够说, 如果 ω 是个 k 次形式, 则 ω 的“拉回”在一个 $(k+1)$ 维区域的 k 维边界上的积分等于 $d\omega$ 的“拉回”在这个 $(k+1)$ 维区域上的积分. 我们不去把这个要求弄得精确, 而只是看一看在 $k = 1$ 而区域是平面上一个三角形的情形. 设 φ 为三角形 $\Delta \subset \mathbb{R}^2$ 的某个邻域到 M 内的一个光滑映射, 并设 Δ 的顶点被 φ 映成了 x, y 和 z(见图 11.5). Δ 的边界由三条曲线 (段)C_1, C_2 和 C_3 组成 (具有适当的定向). 设 ω 为 M 上一个线性微分形式. 我们于是期待着有

$$\int \varphi^* d\omega = \int C_1^* \varphi^* \omega + \int C_2^* \varphi^* \omega + \int C_3^* \varphi^* \omega.$$

如果 $\omega = df$, 则右端的三个积分 (由微积分基本定理) 等于 $f(y) - f(x) + f(z) - f(y) + f(x) - f(z) = 0$. 于是 $\int \varphi^* d(df) = 0$. 因为三角形是任意的, 我们期待有

$$d(df) = 0.$$

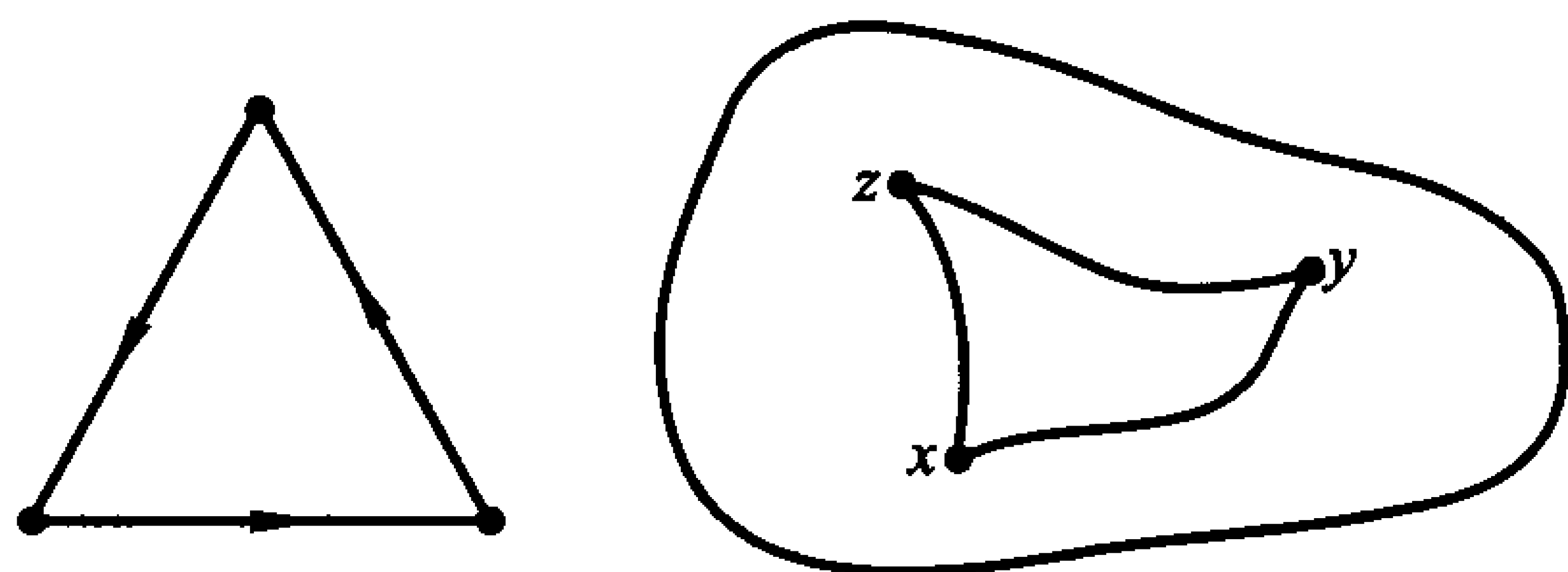


图 11.5

定理 3.2 *存在惟一的线性映射 $d : \wedge^k(M) \longrightarrow \wedge^{k+1}(M)$ 使得在 $\wedge^0$ 上与原来的 d 一致, 并且使得当 $\omega \in \wedge^p(M)$ 时有*

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^p \omega_1 \wedge d\omega_2 \tag{3.12}$$

和当 $f \in \wedge^0(M)$ 时有

$$d(df) = 0. \tag{3.13}$$

证明 首先建立 d 的惟一性. 为此我们看到 (3.1) 式表明 d 是局部的, 这个意思是说如果在某个开集 U 上 $\omega = \omega'$, 则在 U 上有 $d\omega = d\omega'$. 事实上, 设 W 为开集使得 $\overline{W} \subset U$, 又设 φ 为 C^∞ 函数使 $\varphi(x) \equiv 1$ 对 $x \in W$ 成立, 而 $\operatorname{supp}\varphi \subset U$. 于是 $\varphi\omega = \varphi\omega'$ 在 M 上处处相等, 从而 $d(\varphi\omega) = d(\varphi\omega')$. 但由 (3.1) 式, $d(\varphi\omega) = \varphi d\omega + d\varphi \wedge \omega$ 在 W 上等于 $d\omega$, 这是因为在 W 上 $\varphi \equiv 1$, 因而 $d\varphi = 0$. 故 $d\omega = d\omega'$ 在 W 上成立, 但 W 是任意的, 从而在 U 上 $d\omega = d\omega'$.

设 (U, α) 为一分图表, 其坐标为 $x^1, \cdots, x^n$. 每个 $\omega \in \wedge^k(M)$ 可以在 U 上写为

$$\omega = \sum_{i_1 < \cdots < i_k} a_{i_1, \cdots, i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

现在有 [用 (3.1) 和 (3.2) 式, 对 k 进行归纳] $d(dx^{i_1} \wedge \cdots \wedge dx^{i_k}) = 0$. 因此 (3.1) 式表明, 在 U 上有

$$d\omega = \sum da_{i_1, \cdots, i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}. \tag{3.14}$$

(3.3) 式给出了 d 的一个局部公式. 它也表明了 d 的惟一性. 事实上, 我们已经证明了在任意开子集 o 上, 最多只有一个算子 d, 它将 $\wedge^k(o)$ 映到 $\wedge^{k+1}(o)$, 并对 o 满足定理的假定条件. 在集合 $o \cap U$ 上它必定由 (3.3) 式给出.

我们现在断定为了建立 d 的存在性, 只要证明由 (3.3) 式给出的 d [在任意分图表 (U, α) 上] 满足在 $\wedge^k(U)$ 上这个定理的要求. 事实上, 假设我们已经证明了这个结论, 那么, 设 $\mathfrak{a}$ 为 M 的一个总图表, 并对每个分图表 $(U, \alpha) \in \mathfrak{a}$ 用 (3.3) 式定义 $d_\alpha : \wedge^k(U) \longrightarrow \wedge^{k+1}(U)$. 我们想在 U 上令 $d\omega = d_\alpha\omega$. 为了使它能相容, 我们必须证明如果 (W, β) 是另一个分图表则在 $U \cap W$ 上有 $d_\alpha\omega = d_\beta\omega$. 但 d_α 和 d_β 都满足在 $U \cap W$ 上的这个定理的假设, 因此它们必定相等.

因此要证明此定理就只要验证由 (3.3) 式定义的算子 d 作为映射: $\wedge^k(U) \longrightarrow \wedge^{k+1}(U)$ 满足我们的要求即可. 它显然是线性的. 要验证 (3.2) 式, 我们注意到有

$$df = \sum \frac{\partial f}{\partial x^i} dx^i,$$

故

$$\begin{aligned} d(df) &= \sum d\left(\frac{\partial f}{\partial x^i}\right) \wedge dx^i = \sum_{i,j} \left(\frac{\partial^2 f}{\partial x^i \partial x^j} dx^j\right) \wedge dx^i \\ &= \sum_{i<j} \left(\frac{\partial^2 f}{\partial x^j \partial x^i} - \frac{\partial^2 f}{\partial x^i \partial x^j}\right) dx^i \wedge dx^j \\ &= 0 \end{aligned}$$

因为两个混合偏导是相等的.

现转到 (3.1) 式. 因 (3.1) 式的两端对 ω_1, ω_2 分别是线性的；那么只要对 $\omega_1 = adx^{i_1} \wedge \cdots \wedge dx^{i_p}$ 和 $\omega_2 = bdx^{j_1} \wedge \cdots \wedge dx^{j_q}$ 去验证 (3.1) 式即可. 现在 $\omega_1 \wedge \omega_2 = abdx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}$ 而 $d(ab) = bda + adb$；因此,

$$\begin{aligned} d(\omega_1 \wedge \omega_2) =& bda \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q} \\ &+ adb \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}, \end{aligned}$$

而

$$\begin{aligned} d\omega_1 \wedge \omega_2 &= (da \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}) \wedge (bdx^{j_1} \wedge \cdots \wedge dx^{j_q}) \\ &= bda \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}, \\ \omega_1 \wedge d\omega_2 &= (a \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p}) \wedge (db \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}) \\ &= (-1)^p adb \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_p} \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_q}, \end{aligned}$$

于是 (3.1) 式成立. 定理得证. □

从等式 (3.3) 可以得出许多重要的推论.

首先, 对 $\omega \in \wedge^k(M)$ 及任意的 k 可立即得到

$$d(d\omega) = 0. \tag{3.15}$$

(记住, 我们只不过假定了 $k = 0$ 时它成立.)

其次, 设 φ: $M_1 \longrightarrow M_2$ 为可微映射. 则对 $\omega \in \wedge^k(M_2)$ 我们有

$$d\varphi^*\omega = \varphi^* d\omega. \tag{3.16}$$

要验证 (3.5) 式只要证实它对任何一对相容的分图表成立即可. 然而如果 $x^1, \cdots, x^n$ 为 M_2 上的坐标, 那么, 局部地,

$$\omega = \sum a_{i_1,\cdots,i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k},$$

于是我们有

$$\begin{aligned} d\varphi^*\omega &= d\left(\sum \varphi^*(a_{i_1,\cdots,i_k})\varphi^* dx^{i_1} \wedge \cdots \wedge \varphi^* dx^{i_k}\right) \\ &= \sum d\varphi^*(a_{i_1,\cdots,i_k}) \wedge d\varphi^* x^{i_1} \wedge \cdots \wedge d\varphi^* x^{i_k} \\ &= \sum \varphi^* da_{i_1,\cdots,i_k} \wedge d\varphi^* x^{i_1} \wedge \cdots \wedge d\varphi^* x^{i_k} \\ &= \varphi^*\left(\sum da_{i_1,\cdots,i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}\right) \\ &= \varphi^* d\omega. \end{aligned}$$

特别地, 如果 X 是 M 上的一个向量场, 我们得到

$$D_X d\omega = d(D_X\omega). \tag{3.17}$$

习题

3.1 计算下列微分形式的 d.

(a) $\gamma = \sum_1^n (-1)^{i-1} x_i dx_1 \wedge \cdots \wedge dx_{i-1} \wedge dx_{i+1} \wedge \cdots \wedge dx_n$.

(b) $r^{-n}\gamma$, 其中 γ 是 (a) 中的那个, 而 $r = \{x_1^2 + \cdots + x_n^2\}^{\frac{1}{2}}$.

(c) $\sum p_i dq_i$.

(d) $\sin(x^2+y^2+z^2)(xdx+ydy+zdz)$.

设 V 为向量空间, 其上有一个非并的双线性形式和一个定向. 于是我们可以像第七章中那样定义 $*$ 算子. 由于我们把切空间 $T_x(V)$ 等同于 $V, x \in V$, 我们可将 $*$ 算子看作 $\wedge^k(V)$ 到 $\wedge^{n-k}(V)$ 的映射. 例如在 $\mathbb{R}^2$ 中的直角坐标下, 我们有

$$*dx = dy, *dy = -dx, \text{等等}.$$

3.2 证明对 $\mathbb{R}^2$ 中任意函数 f 有

$$d*df = \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right) dx \wedge dy.$$

3.3 用通常的内积得出 $\mathbb{R}^n$ 中关于 $d*d$ 的相似表达式.(回忆有 $*dx^1 \wedge \cdots \wedge dx^k = dx^{k+1} \wedge \cdots \wedge dx^n$, 或更一般有 $*dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \pm dx^{j_1} \wedge \cdots \wedge dx^{j_{n-k}}$, 其中 $(i_1, \cdots, i_k, j_1, \cdots, j_{n-k})$ 是 $(1, \cdots, n)$ 的一个置换, 而 $\pm$ 是此置换的符号.)

3.4 设 x, y, z, t 是 $\mathbb{R}^4$ 的坐标. 在每点的切空间引进内积使得

$$(dx,dx) = (dy,dy) = (dz,dz) = 1,$$
$$(dx,dy) = (dx,dz) = (dx,dt) = (dy,dz) = (dy,dt) = (dz,dt) = 0,$$

而

$$(cdt, cdt) = -1,$$

其中 c 是一个正常数. 设 ω 为一个 2 - 形式, 定义为

$$\omega = c(E_1 dx \wedge dt + E_2 dy \wedge dt + E_3 dz \wedge dt)$$
$$+ B_1 dy \wedge dz + B_2 dz \wedge dx + B_3 dx \wedge dy.$$

又设 γ 为一个 3 - 形式, 定义为

$$\gamma = \rho dx \wedge dy \wedge dz - (J_1 dy \wedge dz + J_2 dz \wedge dx + J_3 dx \wedge dy) \wedge dt.$$

把等式

$$d\omega = 0 \quad \text{和} \quad d * \omega = 4\pi\gamma$$

写为包含各种系数和它们的偏导数的方程式.

11.4 斯托克斯定理

这一节所要证明的一个定理是单变量微积分基本定理的一个大范围的推广. 或许应该把它称作多变量微积分的基本定理. 首先做一些定义.

设 D 是流形 M 中一个具正则边界的区域. 我们回想 (第十章第 6 节)M 的每个点位于三种类型的分图表中的某一个 (U,α) 中.

设 (U,α) 和 (W,β) 是 M 中型 (iii) 的两个分图表. 于是在那里给出了 $J_{\beta\circ\alpha^{-1}}$ 的矩阵为

$$\begin{bmatrix} \dfrac{\partial y^1}{\partial x^1} & \cdots & \dfrac{\partial y^1}{\partial x^{n-1}} & \dfrac{\partial y^1}{\partial x^n} \\ \vdots & & \vdots & \vdots \\ \dfrac{\partial y^{n-1}}{\partial x^1} & \cdots & \dfrac{\partial y^{n-1}}{\partial x^{n-1}} & \dfrac{\partial y^{n-1}}{\partial x^n} \\ 0 & \cdots & 0 & \dfrac{\partial y^n}{\partial x^n} \end{bmatrix},$$

故

$$\begin{aligned} \det J_{\beta\circ\alpha^{-1}} &= \frac{\partial y^n}{\partial x^n} \times \det \begin{bmatrix} \dfrac{\partial y^1}{\partial x^1} & \cdots & \dfrac{\partial y^1}{\partial x^{n-1}} \\ \vdots & & \vdots \\ \dfrac{\partial y^{n-1}}{\partial x^{n-1}} & \cdots & \dfrac{\partial y^{n-1}}{\partial x^{n-1}} \end{bmatrix} \\ &= \frac{\partial y^n}{\partial x^n} \times \det J_{(\beta\restriction\partial D)\circ(\alpha\restriction\partial D)^{-1}}. \end{aligned} \tag{4.1}$$

又由于 $\alpha(U\cap W)\cap\{\boldsymbol{v} : v^n > 0\} = \alpha(U\cap W\cap \mathrm{int}D)$, 故当 $x^n > 0$ 时有 $y^n(x^1, \cdots, x^n) > 0$. 因此在 $x_n = 0$ 上一个边界点有 $\dfrac{\partial y^n}{\partial x^n} > 0$.

现在假设 $\boldsymbol{M}$ 是个定向流形, $D \subset \boldsymbol{M}$ 为具正则边界的区域. 我们要使 ∂D 成为一个定向流形. 我们称一个总图表 $\mathfrak{a}$ 是*已校正的*是说每个 $(U,\alpha) \in \mathfrak{a}$ 是三种类型 (i),(ii),(iii) 之一, 并且 $\mathfrak{a}$ 的每个分图表是正的.

如果 $\dim M > 1$, 我们总能找到已校正的总图表. 事实上, 选取 U 连通, 我们发现每个 (U,α) 或为正或为负. 如果 (U,α) 为负, 我们将其换作 (U,α'), 其中 $x^1_{\alpha'} = -x^1_{\alpha}$.

如果 $\dim M = 1$, 则 ∂D 由一个离散的点集构成 (我们可将其看做“零维流形”). 每个 $x \in \partial D$ 在一个 (iii) 型的分图表中, 它或为正或为负. 我们对任一为负的 (iii) 型分图表 (从而所有缩小了的分图表) 指定给 x 一个正号. 而在 (iii) 型分图表为正时指定给 x 一个负号. 这样我们就如同图 11.6 所展示的那样“定向”了 ∂D.

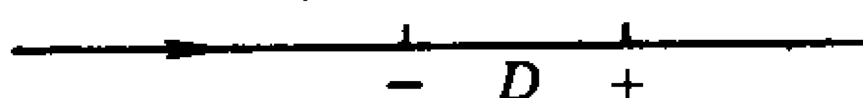


图 11.6

如果 $\dim M > 1$, 我们选取 M 上一个已校正的 (定向) 总图表. 于是由 (4.1) 式和 $\dfrac{\partial y^n}{\partial x^n} > 0$ 得到

$$\det J_{(\beta \restriction \partial D) \circ (\alpha \restriction \partial D)^{-1}} > 0.$$

这表明 $(U \restriction \partial D, \alpha \restriction \partial D)$ 是 ∂D 的一个定向总图表. 因此得到 ∂D 的一个定向. 这不完全是我们所需要的 ∂D 的定向. 我们要在 ∂D 上选取一个定向使得 $(U \restriction \partial D, \alpha \restriction \partial D)$ 具有与 $(-1)^n$ 相同的符号, 其理由很快就能清楚. 这就是说, $(U \restriction \partial D, \alpha \restriction \partial D)$ 当 n 为偶数时是个正分图表, 而在 n 为奇数时我们取定向为由 $(U \restriction \partial D, \alpha \restriction \partial D)$ 所决定的定向相反. 现在可以叙述我们的主要定理了.

定理 4.1 (斯托克斯定理) *设 M 为 n 维的定向流形, $D \subset M$ 是个具正则边界的区域. 设 ∂D 代表作为定向流形的 D 的边界. 于是对任意 $\omega \in \wedge^{n-1}(M)$ 且有紧支集的任意形式, 我们有*

$$\int_{\partial D} \iota^* \omega = \int_D d\omega, \tag{4.2}$$

像通常那样, ι 为边界 ∂D 到 M 的内射.

证明 $n = 1$ 时就是微积分基本定理.

对 $n > 1$ 的证明几乎与第十章的定理 6.1 的证明完全一样. 选取一个已校正的总图表 $\mathfrak{a}$ 和从属于 $\mathfrak{a}$ 的 1 的分解 $\{g_j\}$. 由于 ω 有紧支集, 我们可以写成

$$\omega = \sum g_j \omega,$$

其中取和是有限的. 因为 (4.2) 式的两端是线性的, 故只要对每个取和项 $g_j\omega$ 验证即可. 由于 $\operatorname{supp} g_j\omega \subset U$, 其中 (U, α) 满足 (i),(ii) 或 (iii) 中之一: 这是必须验证的三种可能性.

如果 (U,α) 满足 (i), $\iota^*\omega=0$, 因为 $\operatorname{supp}\omega\cap\partial D=\varnothing$, 又因为 $D\cap\operatorname{supp}\omega=\varnothing$, 故

$$\int_D d\omega=\int_M e_D d\omega=0.$$

因此 (4.2) 式的两端都为零.

如果 (U,α) 满足 (ii) 则 (4.2) 式的左端为 0. 我们必须证明右端也如此. 设 $x^1,\cdots,x^n$ 为 (U,α) 的坐标, 记

$$g_j\omega=a_1dx^2\wedge\cdots\wedge dx^n+a_2dx^1\wedge dx^3\wedge\cdots\wedge dx^n+\cdots+a_ndx^1\wedge\cdots\wedge dx^{n-1}.$$

于是

$$dg_j\omega=\sum(-1)^{i-1}\frac{\partial a_i}{\partial x^i}dx^1\wedge\cdots\wedge dx^n,$$

因而

$$\int dg_j\omega=\sum(-1)^{i-1}\int_{\mathbb{R}^n}\frac{\partial a_i}{\partial x^i}.$$

由于 $g_j\omega$ 有紧支集, 函数 a_i 也具紧支集, 从而在 $\mathbb{R}^n$ 上的积分可以换成在 $\square^{\boldsymbol{R}}_{-\boldsymbol{R}}$ 上的积分, 其中 $\boldsymbol{R}=\prec R,\cdots,R\succ$, 并且 R 选择足够大使得 $\operatorname{supp}a_i\subset\square^{\boldsymbol{R}}_{-\boldsymbol{R}}$. 将此重积分写为一个积分的样子, 我们得到

$$\int_{\square^{\boldsymbol{R}}_{-\boldsymbol{R}}}\frac{\partial a_i}{\partial x^i}=\int_{\mathbb{R}^{n-1}}a_i(\cdots,R,\cdots)-a_i(\cdots,-R,\cdots)=0,$$

这是因为 $a_i(\cdots,R,\cdots)=a_i(\cdots,-R,\cdots)=0$.

现在考察情形 (iii) 时的 $\int dg_j\omega$. 论证完全同于前面一样进行, 但此时要计算的不是 $\int_{\alpha(U)}$ 而是 $\int_{\alpha(U\cap D)}\dfrac{\partial a_i}{\partial x_i}$(见图 11.7).

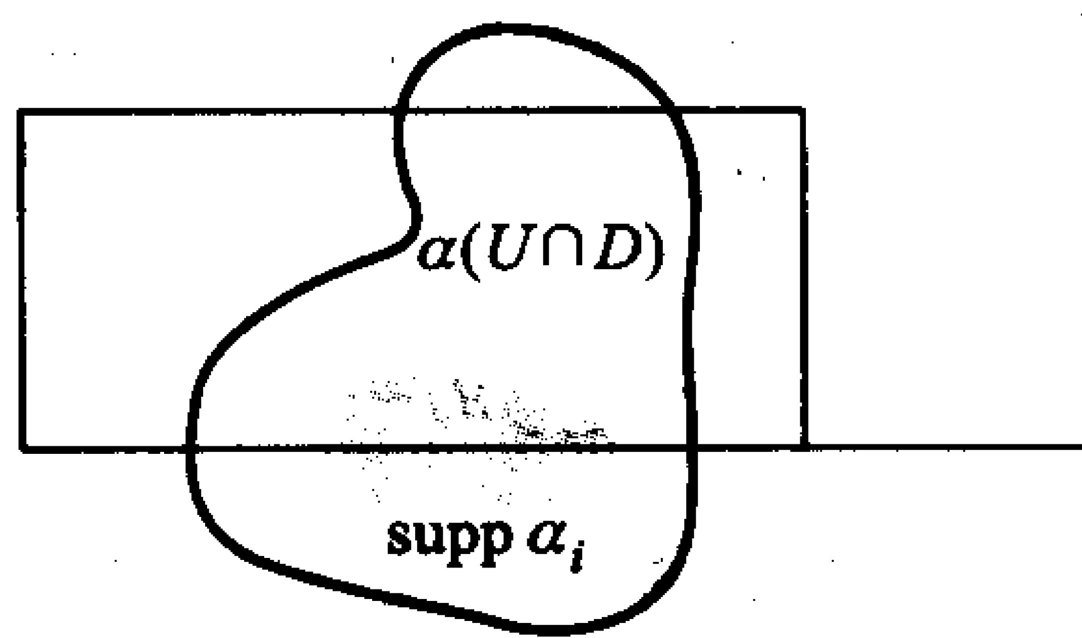


图 11.7

我们现在可将积分区域换作矩形形式 $\square_{\prec -R,\cdots,-R,0\succ}^{\prec R,\cdots,R\succ}$, R 充分大. 如果 $i<n$, 如前面一样有 $\int \dfrac{\partial a_i}{\partial x_i}=0$. 如果 $i=n$, 我们则得到

$$\int_{\partial(U\cap D)} \frac{\partial a_n}{\partial x_n} = -\int_{\mathbb{R}^{n-1}} a_n(\cdot,\cdot,\cdots,\cdot,0).$$

故

$$\int d(g_j\omega)=\sum(-1)^{i-1}\int\frac{\partial a_i}{\partial x^i}=(-1)^n\int_{\mathbb{R}^{n-1}}a_n(\cdot,\cdot,\cdots,\cdot,0).$$

现在因为在 $U\cap\partial D$ 上 $x^n=0$, 我们有 $\iota^*dx^n=0$, 因此

$$\iota^*\omega=(\iota^*a_n)(\iota^*dx^1)\wedge\cdots\wedge(\iota^*dx^{n-1}),$$

或者如果将 $x^1,\cdots,x^{n-1}$ 看作 $(U\upharpoonright\partial D,\alpha\upharpoonright\partial D)$ 的坐标 (不恰当地使用符号), 我们得到

$$\iota^*\omega=a(\cdot,\cdot,\cdots,\cdot,0)dx^1\wedge\cdots\wedge dx^{n-1}.$$

由于我们对 ∂D 所选取的这种定向, 我们最后得到

$$\int_{\partial D}\iota^*\omega=(-1)^n\int_{\mathbb{R}^{n-1}}a_n(\cdot,\cdot,\cdots,\cdot,0).$$

这便完成了定理的证明. □

如同散度定理的情形, 定理 4.1 对我们应用于更一般的区域来说还不够广泛. 为此, 我们再次使用具有几乎正则边界的区域这个概念.

我们已经看到, 具有型 (iii) 邻域的点 $x\in\partial D$ 构成一个微分流形.(我们应该记得, 这些点不必穷尽 ∂D 的所有点.) 相似地, 如果 $\boldsymbol{M}$ 是个定向流形, 则此点集也成为一个定向流形 (依前所述, 以 $(-1)^n$ 乘以诱导定向.) 借用不当言词我们可用 $\partial\boldsymbol{D}$ 表示此定向流形. 因此 $\partial\boldsymbol{D}$ 是个定向流形, 作为集合它却不是 ∂D 而只是 ∂D 的 "正则" 点, 就是说 $\widetilde{\partial D}$ 的点.

定理 4.2 (斯托克斯定理) 设 $\boldsymbol{M}$ 为 n 维定向流形, $D\subset M$ 为具有几乎正则边界的区域. 设 ∂D 如上, ι 为内射 $\partial\boldsymbol{D}\longrightarrow\boldsymbol{M}$. 于是对任意 $\omega\in\wedge^{n-1}(M)$, 具紧支集的形式, 我们有

$$\int_{\partial\boldsymbol{D}}\iota^*\omega=\int_{\boldsymbol{D}}d\omega. \tag{4.3}$$

证明 证明如前. 选取一个已校正的总图表和一个从属于此总图表的 1 的分割. 记 $\omega = \sum g_j\omega$, 而现在要考虑四种情形. 前三种情形已经处理过了. 新的情形是

$$g_j\omega = \sum_j a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n,$$

其中 ^ 表示 dx^j 已被略去, 它的支集在 U 中, 而 (U, α) 是型 (iv) 的一个分图表. 由线性性, 只要对每个右端的每个求和项, 即

$$a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n$$

验证 (4.2) 式即可.

现在除非 $j \geqslant p$, 否则便有 $\iota^*(a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n) = 0$, 其理由是 dx^p 在 $\partial \boldsymbol{D} \cap U$ 上为零, 而 $\partial D \cap U$ 在 α 下的像位于 H_p^k 中. 如果 $j < p$, 则这些 dx^p 出现, 从而 $\iota^*(a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n) = 0$. 如果 $j > p$, 则 $\iota^*(a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n)$ 除了在 ∂D 在 α 下映成了 H_j^k 那部分外其他处处为零.

另一方面,

$$d(a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n) = (-1)^{j-1} \frac{\partial a_j}{\partial x^j} dx^1 \wedge \cdots \wedge dx^n.$$

我们可以在矩形

$$\square_{\prec -R, \cdots, -R, 0, \cdots, 0 \succ}^{\prec R, \cdots, R, \cdots, R \succ}$$

上积分来计算积分 $\int_D$(其中这些 $-R$ 延伸超过了第 $(k-1)$ 个位置). 首先对 x^j 积分我们得到

$$\int_D d(a_j dx^1 \wedge \cdots \wedge \widehat{dx^j} \wedge \cdots \wedge dx^n) = (-1)^j \int_{H_i^k} a_j.$$

另一方面, H_j^k 的定向使此积分有了必须的符号令 (4.2) 式成立. 定理 4.2 得证. □

如同前面那样, 我们可以应用定理 4.1 和 4.2 到更为一般的区域, 办法是用极限论证. 例如, 这里所叙述的定理 4.2 不能用于图 11.8 的区域 D, 因为此时的曲线 C_1 和 C_2 在 P 点相切. 然而用 "分割成小段" 的办法可以应用于这样得到的近似区域 (图 11.9), 并且显然 (4.2) 式两端对 D' 的值接近于对 D 的值. 那么

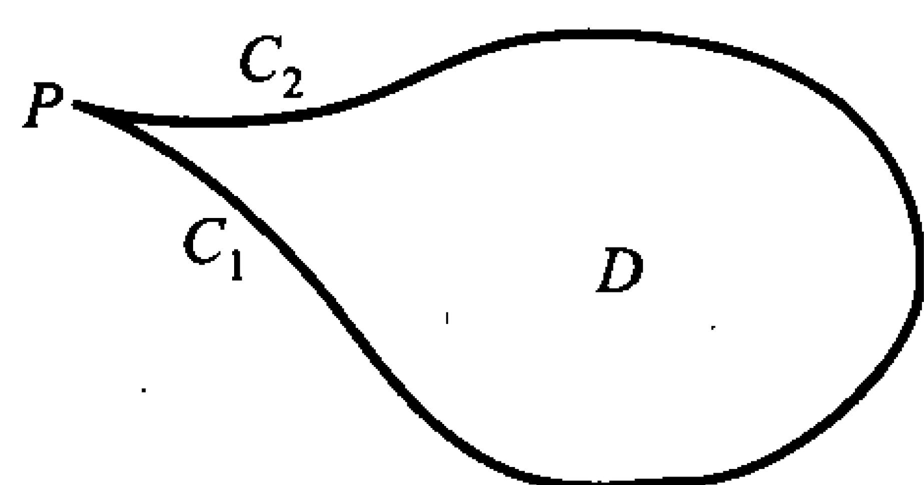


图 11.8

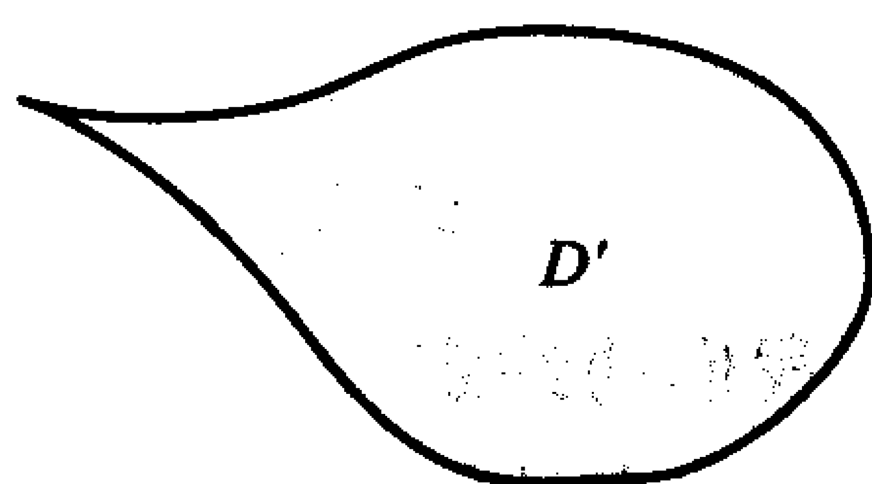


图 11.9

通过极限可得到 (4.2) 式对 D 的值. 仍像从前那样, 我们不去叙述包括这些情形的更一般的定理. 在每种情形该如何应用极限推理是清楚的.

由于斯托克斯定理的陈述和证明与散度定理的是如此相近, 读者会推测是否其中一个蕴涵了另一个. 在定向流形上, 散度定理的确是斯托克斯定理的一个推论. 要明白这点, 设 Ω 是 $\wedge^n(M)$ 中对应于密度 ρ 的一个元素. 如果 X 是个向量场, 则可清楚看到, n - 形式 $D_X\Omega$ 对应于密度 $D_X\rho = \operatorname{div} \prec X, \rho \succ$. 提前使用将在第 6 节引进的记号: 设 $X\rfloor\Omega$ 为一个 $(n-1)$ - 形式, 其定义为

$$X\rfloor\Omega(\xi^1, \cdots, \xi^{n-1}) = (-1)^{n-1}\Omega(\xi^1, \cdots, \xi^{n-1}, X).$$

借助于坐标, 如果 $\Omega = a dx^1 \wedge \cdots \wedge dx^n$, 则

$$\begin{aligned} X\rfloor\Omega = & a[X^1 dx^2 \wedge \cdots \wedge dx^n - X^2 dx^1 \wedge dx^3 \wedge \cdots \wedge dx^n + \cdots \\ & + (-1)^{n-1} X^n dx^1 \wedge \cdots \wedge dx^{n-1}]. \end{aligned}$$

注意,

$$d(X\rfloor\Omega) = \left(\sum \frac{\partial a X^i}{\partial x^i}\right) dx^1 \wedge \cdots \wedge dx^n,$$

它对应于密度 $D_X\rho = \operatorname{div} \prec X, \rho \succ$, 故其正是 n - 形式 $D_X\Omega$. 于是由斯托克斯定理表明

$$\int_{\partial D} \iota^*(X\rfloor\Omega) = \int_D d(X\rfloor\Omega) = \int_D \operatorname{div} \prec X, \rho \succ.$$

我们必须把 $X\rfloor\Omega$ 与在 ∂D 上的密度 ρ_X 进行比较. 由 (2.2) 式, 它们在所有地方除符号外都是相同的. 要验证符号也相同, 只要比较在任意 $x \in \partial D$ 处

$$\rho_X\left(\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^{n-1}}\right) = \rho\left(\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^{n-1}}, X\right)$$

与

$$\iota^*\left(X\rfloor\Omega\left(\frac{\partial}{\partial x^1}, \cdots, \frac{\partial}{\partial x^{n-1}}\right)\right)$$

即可. 现在

$$\iota^*(X\rfloor\Omega) = (-1)^{n-1}X^n dx^1 \wedge \cdots \wedge dx^{n-1},$$

而且根据我们的约定, $x^1, \cdots, x^{n-1}$ 是依 $(-1)^n$ 决定是个正的还是负的坐标系. 因此两者一致当且仅当 X^n 为负, 即

$$\int_{\partial D} \iota^*(X\rfloor\Omega) = \int_{\partial D} \varepsilon_X \rho_X.$$

习题

4.1 直接地同时也使用斯托克斯定理计算下列曲面积分. 设 □ 表示单位立方体, B 为 $\mathbb{R}^3$ 中的单位球.

(a) $\int_{\partial\square} xdy \wedge dz + ydz \wedge dx + zdx \wedge dy$,

(b) $\int_{\partial B} x^3 dy \wedge dz$,

(c) $\int_{\partial\square} \cos z dx \wedge dy$,

(d) $\int_{\partial U} xdy \wedge dz$, 其中 $U = \{(x,y,z) : x \geqslant 0, y \geqslant 0, z \geqslant 0, x^2+y^2+z^2 \leqslant 1\}$.

4.2 设 $\omega = yzdx + xdy + dz$. γ 为平面上反时针定向的单位圆. 计算 $\int_\gamma \omega$. 设

$$A_1 = \{(x,y,z) : z = 0, x^2+y^2 \leqslant 1\},$$
$$A_2 = \{(x,y,z) : z = 1-x^2-y^2, x^2+y^2 \leqslant 1\}.$$

定向曲面 A_1 和 A_2 使得 $\partial \boldsymbol{A}_1 = \partial \boldsymbol{A}_2 = \gamma$. 通过做积分验证 $\int_{\boldsymbol{A}_1} d\omega = \int_{\boldsymbol{A}_2} d\omega = \int \omega$.

4.3 设 S^1 是圆, $\omega = \left(\dfrac{1}{2\pi}\right) d\theta$, 其中 θ 为角坐标.

(a) 设 $\varphi : S^1 \longrightarrow S^1$ 为可微映射. 证明 $\int \varphi^*\omega$ 是个整数. 称此整数为 φ 的次数, 以 $\deg\varphi$ 表示.

(b) 设 φ_t 是一组映射 (每个 t 有一个), 它可微地依赖于 t. 证明 $\deg\varphi_0 = \deg\varphi_1$.

(c) 设我们把 S^1 看作为在复数中的单位圆. 设 f 为在复数上的函数, 并设对 $|z| = r$ 时 $f(z) \neq 0$. 由令 $\varphi_{r,f}(e^{i\theta}) = \dfrac{f(re^{i\theta})}{|f(re^{i\theta})|}$ 定义的 $\varphi_{r,f}$. 假设 $f(z) = z^n$. 计算 $\deg\varphi_{r,f}, r \neq 0$.

(d) 设 f 为次数 $n \geqslant 1$ 的多项式. 于是

$$f(z) = a_n z^n + a_{n-1}z^{n-1} + \cdots + a_0,$$

其中 $a_n \neq 0$. 证明至少存在一个复数 z_0 使 $f(z_0) = 0$.

[提示: 设若相反. 则 $\varphi_{r,\left(\frac{1}{a_n}\right)f}$ 对 $0 \leqslant r < \infty$ 有定义, 而由 (b), $\deg\varphi_{r,\left(\frac{1}{a_n}\right)f} =$ 常值. 计算此表达式的 $\lim_{r=0}$ 和 $\lim_{r=\infty}$.] 设 X 为定义于 $\mathbb{E}^2$ 中原点的某个邻域 U

中的向量场, 并设 $X(0)=0$ 而 $X(x)\neq 0$, 当 $x\neq 0$. 于是 X 只在原点为 0. 定义映射 $\varphi_r: S^1 \longrightarrow S^1$ 为

$$\varphi_r(e^{i\theta}) = \frac{X(re^{i\theta})}{\|X(re^{i\theta})\|}.$$

这个映射对充分小的 r 是有定义的. 由习题 4.3(b), 此映射的次数不依赖于 r. 这个次数被称作此向量场 X 在原点的指数.

4.4 计算下面向量场的指数.

(a) $x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial x}$.

(b) $x\dfrac{\partial}{\partial x} - y\dfrac{\partial}{\partial y}$.

(c) $y\dfrac{\partial}{\partial x} - x\dfrac{\partial}{\partial y}$.

(d) 构造一个指数为 2 的向量场.

(e) 证明对任意向量场 $X, -X$ 的指数与 X 的指数相同.

4.5 设 X 为一个定向二维流形 M 上的向量场, 假定对某个 $p\in M$ 有 $X(p)=0$, 而在 p 的一个小邻域的任意其他的点上 X 不为零. 选取一个定向的分图表, 它把 p 映到原点, 我们从而得到一个 $\mathbb{E}^2$ 中的向量场, 它在原点为 0. 证明这个向量场的指数不依赖于分图表的选取. 因此可定义 X 在 p 的指数.

4.6 (a) 在球面 S^2 上设 X 为一个向量场, 它处处切于经圆且只在南北极为零. 在每个极点它的指数是什么?

(b) Y 是个向量场, 它处处切于纬圆, 而只在南北极为零. 在每个极点的指数是多少?

11.5 斯托克斯定理的一些例示

我们叙述定理 4.2 的一个简单而重要的推论.

定理 5.1 设 $\varphi: M_1 \to M_2$ 为从 k 维定向流形 M_1 到 n 维流形 M_2 的一个可微映射. 设 ω 为 M_2 上一个 $k-1$ 次形式, $D\subset M_1$ 为具有在 M_1 上几乎正则边界的区域. 于是我们有

$$\int_{\partial D} \iota^*\varphi^*\omega = \int_D \varphi^*(d\omega). \tag{5.1}$$

等式 (5.1) 直接从 (4.2) 式和 $\varphi^* d = d\varphi^*$ 这个事实得到.

我们可将 (5.1) 式的右端看作 $d\omega$ 在"定向 k 维超曲面" $\varphi(\boldsymbol{D})$ 上的积分. 等式 (5.1) 说这个积分等于 ω 在 $(k-1)$ 维超曲面 $\varphi(\partial\boldsymbol{D})$ 上的积分.

我们现在给出定理 5.1 的一个简单应用. 设 $C_0:[0,1]\to M$, $C_1:[0,1]\to M$ 为两条可微曲线, $C_0(0)=C_1(0)=p$ 和 $C_0(1)=C_1(1)=q$ (见图 11.10). 我们说 C_0 和 C_1 为 (微分) 同伦是指存在一个单位正方形 $[0,1]\times[0,1]\subset\mathbb{R}^2$ 到 M 中的可微映射 φ 使得 $\varphi(t,0)=C_0(t),\varphi(t,1)=C_1(t),\varphi(0,s)=p,\varphi(1,s)=q$ (见图 11.11). 对 s 的每个值我们得到由 $C_s(t)=\varphi(t,s)$ 的曲线 C_s. 可以想像 φ 为提供了从曲线 C_0 到曲线 C_1 的一个可微"形变".

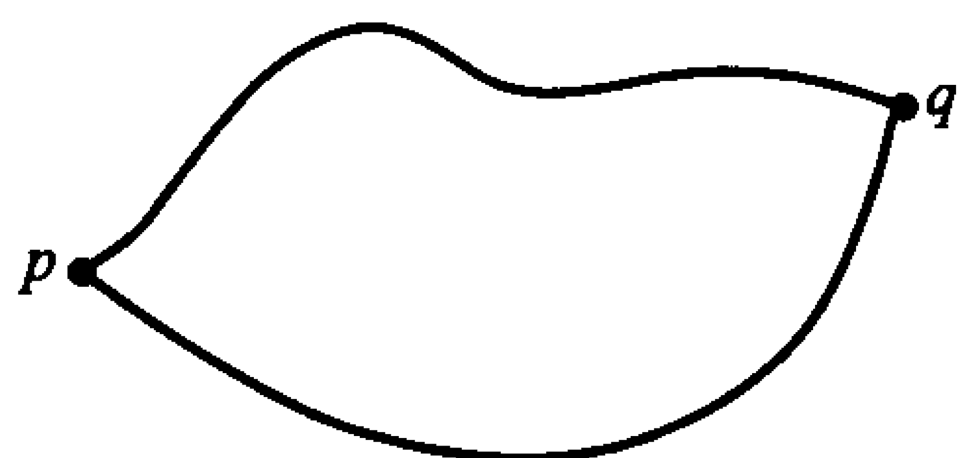
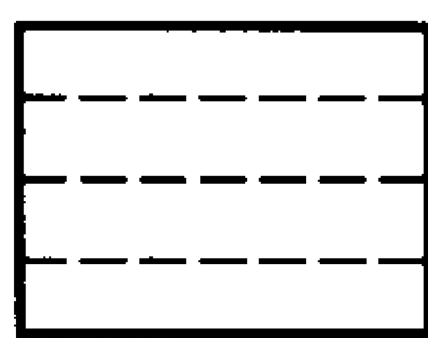


图 11.10



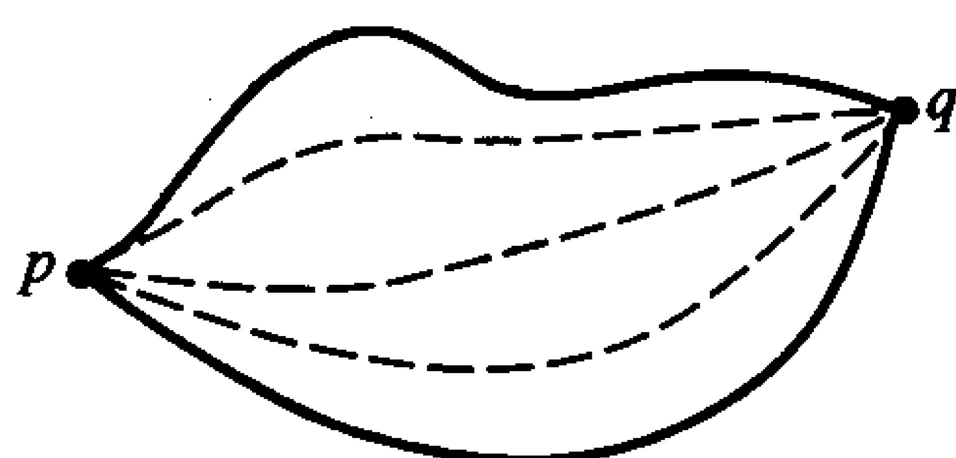


图 11.11

命题 5.1 设 C_0 和 C_1 为可微同伦曲线, 且设 ω 是 M 上的线性微分形式, 满足 $d\omega=0$. 于是

$$\int_{C_1}\omega=\int_{C_2}\omega. \tag{5.2}$$

证明 事实上,

$$\int_{\partial\square^{\prec 1,1\succ}_{\prec 0,0\succ}}\varphi^*\omega=\int_{\square^{\prec 1,1\succ}_{\prec 0,0\succ}}\varphi^*d\omega=0.$$

但 $\int_{\partial\square}$ 是对应于正方形四条边的四项之和. 两个垂直边 ($t=0$ 和 $t=1$) 对此积分无贡献, 这是因为 φ 将曲线映射成了点. 顶边给出了 $-\int_{C_1}$(因为逆时针方向), 而底边给出 $\int_{C_0}$. 因此 $\int_{C_0}\omega-\int_{C_1}\omega=0$, 命题得证. □

容易看出此命题可毫无困难地拓展到逐段可微的曲线和逐段可微的同伦. 我们说两条逐段可微曲线 C_0 和 C_1 都是 (逐段可微) 同伦的, 如果存在连续映射 $\varphi:[0,1]\times[0,1]\to M$ 使得

(i) $\varphi(0,s)=p,\varphi(1,s)=q$;

(ii) $\varphi(t,0)=C_0(t),\varphi(t,1)=C_1(t)$;

(iii) 存在有限个点 $t_0<t_1<\cdots<t_m$ 使得 φ 与定义在每个矩形 $[t_i,t_{i+1}]\times[0,1]$ 的某一邻域由可微映射的限制是一致的 (见图 11.12).

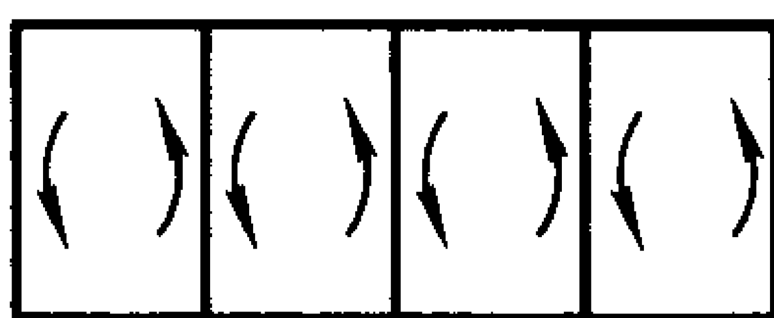


图 11.12

要验证命题 5.1 对于逐段可微同伦成立, 我们把斯托克斯定理用于每个矩形并看到那些内部垂直线所做的贡献都彼此消去了.

我们称一个流形 M 是连通的是指每一对点可以由一条 (逐段可微的) 曲线相联结. 因此 $\mathbb{R}^n$ 是个连通的例子. 我们称 M 是单连通的是说所有 (逐段可微的) 联结相同两个点的曲线都是 (逐段可微) 同伦的 (注意圆, S' 不是单连通的). 我们来验证 $\mathbb{R}^n$ 是单连通的. 如果 C_0 和 C_1 为两条曲线, 设 $\varphi : [0,1]\times[0,1]\to\mathbb{R}^n$ 由

$$\varphi(t,s) = sC_0(t) + (1-s)C_1(t)$$

给出, 显然 φ 具有所要的性质.

命题 5.2 设 M 为连通的单连通流形, 设 $O\in M$ 及 $\omega\in\wedge^1(M)$ 满足 $d\omega=0$. 对任意 $x\in M$ 令 $f(x)=\int_C\omega$, 其中 C 为联结 O 到 x 的某条逐段可微的曲线. 于是, 函数 f 是有明确定义和可微的, 并且 $df=\omega$.

证明 由命题 5.1 知 f 是有明确定义的. 如果 C_0 和 C_1 为两条联结 O 到 x 的曲线, 则它们同伦, 故 $\int_{C_0}\omega=\int_{C_1}\omega$. 显然 f 是连续的, 这是因为

$$f(x)-f(y)=\int_D\omega,$$

其中 D 为联结 y 到 x 的任意曲线 (图 11.13).

要验证 f 可微, 设 (U,α) 是 x 的一个分图表, 坐标为 $\prec x^1,\cdots,x^n\succ$. 于是

$$f(x^1,\cdots,x^i+h,\cdots,x^n)-f(x^1,\cdots,x^n)=\int_C\omega,$$

其中 C 是联结 p 到 q 的任意曲线, 而 $\alpha(p)=(x^1,\cdots,x^i,\cdots,x^n),\alpha(q)=(x^1,\cdots,x^i+h,\cdots,x^n)$. 可取 C 为

$$\alpha\circ C(t)=(x^1,\cdots,x^i+ht,\cdots,x^n)$$

给出的曲线. 如果 $\omega = a_1dx^1 + \cdots + a_ndx^n$, 则

$$\int_C \omega = \int_0^1 ha_idt = \int_0^h a_i(x^1, \cdots, x^i + s, \cdots, x^n)ds$$

(见图 11.14). 因此

$$\lim_{h\to 0}\frac{1}{h}[f(x^1, \cdots, x^i + h, \cdots, x^n) - f(x^1, \cdots, x^n)] = a^i,$$

那就是说 $\frac{\partial f}{\partial x^i} = a^i$. 这证明了 f 是可微的, 并且 $df = \omega$, 从而命题得证. □

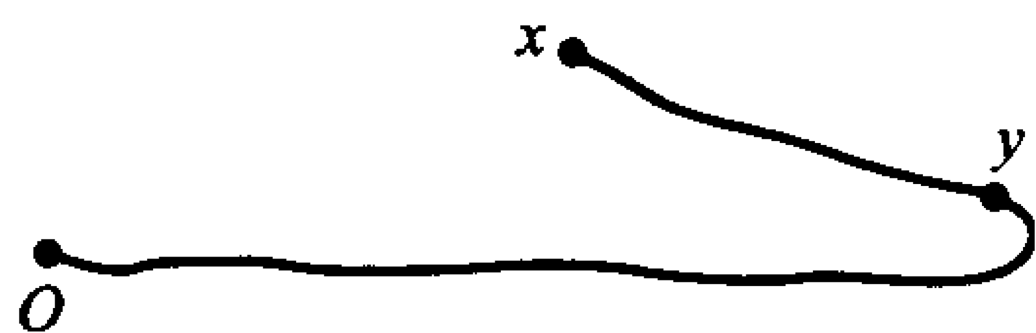


图 11.13

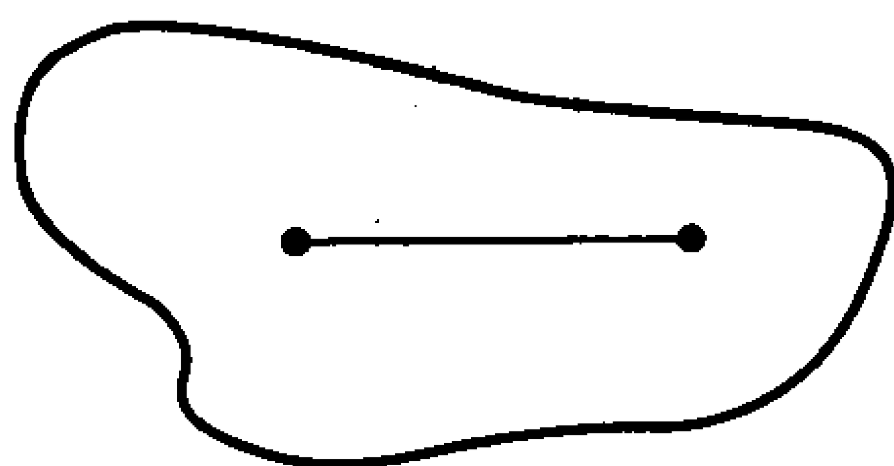


图 11.14

我们已经得到了结论, 即每个满足 $d\omega = 0$ 的 $\omega \in \wedge^1(\mathbb{R}^n)$ 具有形式 df. 更一般地, 可以建立这样的结论, 即如果 $\Omega \in \wedge^k(\mathbb{R}^n)$ 满足 $d\Omega = 0$, 则有某个 $\omega \in \wedge^{k-1}(\mathbb{R}^n)$ 使得 $\Omega = d\omega$.

* 对任意流形而言, 这是不对的. 例如, 每个 $\omega \in \wedge^1(S^1)$ 满足 $d\omega = 0$. 然而角形式的元素 (它不恰当地被记作 $d\theta$) 不是任何函数的 d. $d^2 = 0$ 表明, 如果 $\Omega = d\omega$, 则 $d\Omega = 0$. 因此空间 $d[\wedge^{k-1}(M)] \subset \wedge^k(M)$ 是空间 $\ker_k d$ 的子空间, 后者是 $\wedge^k(M)$ 中满足 $d\Omega = 0$ 的元素构成的空间. 商空间 $\ker_k d/d[\wedge^{k-1}]$ 被记作 $H^k(M)$, 称其为 M 的第 k 个上同调群. 如果 M 是紧的, 则可证明 H^k 是有限维的. 它 (粗略地说) 度量了在 M 中"有多少" k 维洞. *

11.6 微分形式的李导数

设 M 为微分流形, φ 为 M 上一个流, 其无穷小生成元为 X. 对任意 $\omega \in \wedge^k(M)$, 我们可以考虑表达式

$$\frac{\varphi_t^*\omega - \omega}{t}.$$

不难验证 (用局部表达式) 当 $t \to 0$ 时这个极限存在并还是 $\wedge^k(M)$ 中的元素, 我们以 $D_X\omega$ 表示. 这节的目的是要提供一个计算 $D_X\omega$ 的有效公式. 为此, 我们先

汇集 D_X 的某些性质. 首先, 我们知道它是线性的:

$$D_X(\omega_1+\omega_2)=D_X\omega_1+D_X\omega_2. \tag{6.1}$$

其次, 我们有

$$\begin{aligned}\varphi_t^*(\omega_1\wedge\omega_2)-\omega_1\wedge\omega_2&=(\varphi_t^*\omega_1)\wedge(\varphi_t^*\omega_2)-\omega_1\wedge\omega_2\\&=(\varphi_t^*\omega_1)\wedge(\varphi_t^*\omega_2)-(\varphi_t^*\omega_1)\wedge\omega_2+(\varphi_t^*\omega_1)\wedge\omega_2-\omega_1\wedge\omega_2.\end{aligned}$$

除以 t 并取极限, 我们看到

$$D_X(\omega_1\wedge\omega_2)=(D_X\omega_1)\wedge\omega_2+\omega_1\wedge(D_X\omega_2). \tag{6.2}$$

最后, 由于 $\varphi_t^*d=d\varphi_t^*$, 我们有

$$D_Xd\omega=d(D_X\omega). \tag{6.3}$$

实际上, 这三个公式对于计算而言已相当充分了. 如果

$$\omega=\sum a_{i_1,\cdots,i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k},$$

则

$$\begin{aligned}D_X\omega=&\sum D_X(a_{i_1,\cdots,i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k}) && \text{由(6.1)式}\\=&\sum[(D_Xa_{i_1,\cdots,i_k})dx^{i_1}\wedge\cdots\wedge dx^{i_k}+a_{i_1,\cdots,i_k}(D_Xdx^{i_1})\wedge\cdots\wedge dx^{i_k}\\&+\cdots+a_{i_1,\cdots,i_k}dx^{i_1}\wedge\cdots\wedge(D_Xdx^{i_k})] && \text{反复使用 (6.2) 式}\\=&\sum[(D_Xa_{i_1,\cdots,i_k})dx^{i_1}\wedge\cdots\wedge dx^{i_k}+a_{i_1,\cdots,i_k}d(D_Xx^{i_1})\wedge\cdots\wedge dx^{i_k}\\&+\cdots+a_{i_1,\cdots,i_k}dx^{i_1}\wedge\cdots\wedge d(D_Xx^{i_k})] && \text{由(6.3)式.}\end{aligned}$$

由于这个表达式相当麻烦 ($d(D_Xx^i)$ 必须展开并集项), 我们要导出一个较简单且更方便的 $D_X\omega$ 的表达式. 为此我们到代数里去走一圈.

回想一下算子 $d:\wedge^k(M)\longrightarrow\wedge^{k+1}(M)$; 它是线性的并满足等式

$$d(\omega_1\wedge\omega_2)=d\omega_1\wedge\omega_2+(-1)^k\omega_1\wedge d\omega_2, \tag{6.4}$$

其中 $\omega_1\in\wedge^k(M)$. 更一般地, 任一映射 (一系列线性的映射) $\theta:\wedge^k(M)\longrightarrow\wedge^{k+1}(M)$ 如果它们满足等式

$$\theta(\omega_1\wedge\omega_2)=\theta\omega_1\wedge\omega_2+(-1)^k\omega_1\wedge\theta\omega_2 \tag{6.4'}$$

和条件

$$\operatorname{supp}\theta(\omega)\subset\operatorname{supp}\omega^1,\tag{6.5}$$

则称其为代数 $\wedge(M)$ 的一个反导数.

由 (6.5) 式得到, 如果在一个开集 U 上, $\omega_1\equiv\omega_2$, 则在 U 上, $\theta(\omega_1)\equiv\theta(\omega_2)$. 现在对每个 $x\in M$, 我们可以找到邻域 U 及函数 $x^1,\cdots,x^n$, 使得 $\omega\in\wedge^k(M)$ 可以在 U 上写成为

$$\omega\equiv\sum a_{i_1,\cdots,i_k}dx^{i_1}\wedge\cdots\wedge dx^{i_k}.\tag{6.6}$$

于是重复运用 (6.4′) 式我们便有

$$\begin{aligned}\theta(\omega)\equiv\sum[&\theta(a_{i_1,\cdots,i_k})\wedge dx^{i_1}\wedge\cdots\wedge dx^{i_k}+a_{i_1,\cdots,i_k}\theta(dx^{i_1}\wedge\cdots\wedge dx^{i_k})\\&+\cdots+(-1)^{k-1}a_{i_1,\cdots,i_k}dx^{i_1}\wedge\cdots\wedge\theta(dx^{i_k})].\end{aligned}\tag{6.7}$$

我们因而得到一个重要结论:

命题 6.1 *任一反导数 $\theta:\wedge^k(M)\longrightarrow\wedge^{k+1}(M),k=0,\cdots,n$, 被它在 $\wedge^0(M)$ 和 $\wedge^1(M)$ 上的作用惟一决定. 就是说, 如果对所有 $\omega\in\wedge^0(M)$ 和 $\wedge^1(M)$ 有 $\theta_1(\omega)=\theta_2(\omega)$, 则对所有 $\Omega\in\wedge^k(M)$ 和所有 k, $\theta_1(\Omega)=\theta_2(\Omega)$.*

现在假设我们已知映射

$$\theta:\wedge^0(M)\longrightarrow\wedge^1(M)\quad 和\quad\theta:\wedge^1(M)\longrightarrow\wedge^2(M),$$

满足 (6.5) 和 (6.4′) 式, 即在它有明确意义的地方, 成立

$$\theta(f_1f_2)=\theta(f_1)f_2+f_1\theta(f_2)\quad 和\quad\theta(f\omega)=\theta(f)\wedge\omega+f\theta(\omega).\tag{6.8}$$

于是, 由 (6.7) 式, 任意分图表 (U,α) 定义了 $\theta:\wedge^k(U)\longrightarrow\wedge^{k+1}(U)$. 这给出了 U 上的一个反导数 θ_U, 它可以用定理 3.1 的证明中的推理很容易地得到验证. 由惟一性论证知, 如果 (W,β) 是第二个分图表, 则反导数 θ_U 和 θ_W 在 $U\cap W$ 上重合. 因此等式 (6.7) 是相容的, 从而产生了 $\wedge(M)$ 上一个有明确定义的反导数.(注意, 对于这个反导数的更一般的情形, 我们只重复了定理 3.1 证明的大约三分之二.)

---

[1]这个条件实际上是 (6.4) 式的一个推论. 事实上, 设 U 为包含 $\operatorname{supp}\omega$ 的开集. 由于 $\{U,M-\operatorname{supp}\omega\}$ 为 M 的开覆盖, 故可找到从属于它的一个 1 的分解. 特别, 可以找到一个 C^∞ 函数 φ, 它在 $\operatorname{supp}\omega$ 上恒等于 1 而在 U 外为 0. 故 $\omega=\varphi\omega$, 从而 $\theta(\omega)=\theta(\varphi\omega)=\theta(\varphi)\wedge\omega+\varphi\theta(\omega)$. 于是, $\operatorname{supp}\theta(\omega)\subset\operatorname{supp}\omega\cup\operatorname{supp}\varphi\subset U$. 因为 U 是 $\operatorname{supp}\omega$ 的任意邻域, 我们得到 $\operatorname{supp}\theta(\omega)\subset\operatorname{supp}\omega$.

我们在上面的论证中也看出, 如果把 $\theta : \wedge^k(M) \longrightarrow \wedge^{k+1}(M)$ 换成 $\theta : \wedge^k(M) \longrightarrow \wedge^{k-1}(M)$ 则没有什么要改变.[我们对于 $f \in \wedge^0(M)$,θ 的作用表示 $\theta(f) = 0$.] 事实上, 同样的推理对任意奇整数 r,

$$\theta : \wedge^k \longrightarrow \wedge^{k+r}$$

也一样行得通. 因此我们可以叙述下面的

命题 6.2 设 $\theta : \wedge^0(M) \longrightarrow \wedge^r(M)$ 及 $\theta : \wedge^1(M) \longrightarrow \wedge^{r+1}(M)$ 为满足 (6.5) 和 (6.8) 式的线性映射, 其中 r 为奇数. 于是存在一个且惟一的一个方式将 θ 扩展为满足 (6.4) 式的反导数 $\theta : \wedge^k(M) \longrightarrow \wedge^{k+r}(M)$.

作为此命题的应用, 我们要把一个反导数 $\theta(X) : \wedge^k(M) \longrightarrow \wedge^{k-1}(M)$ 联系到 M 上每个光滑向量场. 因为 $r = -1$, 对 $f \in \wedge^0(M)$, 令

$$\theta(X)f = 0.$$

对 $\omega \in \wedge^1(M)$ 令

$$\theta(X)\omega = \prec X, \omega \succ . \tag{6.9}$$

验证 (6.8) 式意味着验证

$$\theta(X)(f\omega) = f\theta(X)\omega,$$

即

$$(X, f\omega) = f \prec X, \omega \succ,$$

这是显然的.

如果 f 为函数, θ 为一反导数, 我们以 $f\theta$ 记 $\omega \longmapsto f\theta(\omega)$ 映射. 容易验证这又是一个反导数.

作为惟一性定理的推论我们可作如下断言: 设 X 和 Y 为光滑向量场, f 和 g 为光滑函数. 则

$$\theta(fX + gY) = f\theta(X) + g\theta(Y). \tag{6.10}$$

由上面的命题, 只要对所有的 $\omega \in \wedge^1(M)$ 验证 (6.10) 式即可. 由 (6.9) 式, 这正好是

$$\prec fX + gY, \omega \succ = f \prec X, \omega \succ + g \prec Y, \omega \succ ,$$

这是显然的.

特别在一个分图表 (U,h) 中, 如果

$$X = X^1\frac{\partial}{\partial x^1} + \cdots + X^n\frac{\partial}{\partial x^n},$$

则

$$\theta(X) = \sum X^i\theta\left(\frac{\partial}{\partial x_i}\right).$$

要计算 $\theta\left(\frac{\partial}{\partial x^i}\right)$, 我们运用 (6.8) 式及

$$\theta\left(\frac{\partial}{\partial x^i}\right)f = 0, \theta\left(\frac{\partial}{\partial x^i}\right)dx^j = \begin{cases} 0, & \text{如果 } i \neq j \\ 1, & \text{如果 } i = j. \end{cases}$$

这个事实. 因此, 举例来说, $\theta\left(\frac{\partial}{\partial x^i}\right)dx^p \wedge dx^q$ 在 $p \neq i$, 或 $q \neq i$ 时为 0, 而 $\theta\left(\frac{\partial}{\partial x^i}\right)dx^i \wedge dx^j = dx^j, \theta\left(\frac{\partial}{\partial x^i}\right)(dx^j \wedge dx^i) = -dx^j$, 等等.

我们称当 s 为偶数时的 (一系列) 映射 $D : \wedge^k(M) \longrightarrow \wedge^{k+s}(M)$ 为一个导数是说它满足 (6.5) 式以及

$$D(\omega_1 \wedge \omega_2) = D\omega_1 \wedge \omega_2 + \omega_1 \wedge D\omega_2. \tag{6.11}$$

因为 s 是偶数, 故这是相容的. 最重要的例子是 D_X, 这时 $s = 0$. 于是 (6.11) 式正好是 (6.2) 式.

所有前面关于扩张的存在性与惟一性的论证可以一成不变地应用于导数, 这极易验证. 因此我们断言:

命题 6.3 设 $D : \wedge^0(M) \longrightarrow \wedge^{0+s}(M)$ 和 $D : \wedge^1(M) \longrightarrow \wedge^{1+s}(M)$ 为满足 (6.5) 和 (6.8) 式的映射 (将 θ 换作 D), 其中 s 为偶数. 于是存在一个且惟一的一个方式把 D 扩展为 $\wedge(M)$ 上的导数.

我们需要一个进一步的代数结果.

命题 6.4 设 $\theta_1 : \wedge^k \longrightarrow \wedge^{k+r_1}, \theta_2 : \wedge^k \longrightarrow \wedge^{k+r_2}$ 为反导数. 则 $\theta_1\theta_2 + \theta_2\theta_1 : \wedge^k \longrightarrow \wedge^{k+r_1+r_2}$ 是一个导数.

证明 因为 r_1 和 r_2 均为奇数, 故 r_1+r_2 为偶. 等式 (6.5) 显然成立. 要验证 (6.4′) 式, 我们设 $\omega_1\in\wedge^k(M)$. 于是

$$\begin{aligned}\theta_1\theta_2(\omega_1\wedge\omega_2)&=\theta_1[\theta_2\omega_1\wedge\omega_2+(-1)^k\omega_1\wedge\theta_2\omega_2]\\&=\theta_1\theta_2\omega_1\wedge\omega_2+(-1)^{k+r_2}\theta_2\omega_1\wedge\theta_1\omega_2\\&\quad+(-1)^k\theta_1\omega_1\wedge\theta_2\omega_2+\omega_1\wedge\theta_1\theta_2\omega_2,\end{aligned}$$

相似地,

$$\begin{aligned}\theta_2\theta_1(\omega_1\wedge\omega_2)&=\theta_2\theta_1\omega_1\wedge\omega_2+(-1)^{k+r_1}\theta_1\omega_1\wedge\theta_2\omega_2\\&\quad+(-1)^k\theta_2\omega_1\wedge\theta_1\omega_2+\omega_1\wedge\theta_2\theta_1\omega_2.\end{aligned}$$

因为 r_1 和 r_2 都是奇数, 所以当我们相加时中间的项相互消去. 从而我们得到了

$$(\theta_1\theta_2+\theta_2\theta_1)(\omega_1\wedge\omega_2)=(\theta_1\theta_2+\theta_2\theta_1)\omega_1\wedge\omega_2+\omega_1\wedge(\theta_1\theta_2+\theta_2\theta_1)\omega_2.\qquad\square$$

作为命题 6.3 的第一个应用, 我们注意到有

$$\theta(X)\circ\theta(Y)=-\theta(Y)\circ\theta(X).\tag{6.12}$$

事实上, 由命题 6.4, $\theta(X)\theta(Y)+\theta(Y)\theta(X)$ 是次数为 -2 的导数, 即它在 $\wedge^0$ 和 $\wedge^1$ 上为零, 从而必定恒为零. 当然, 我们也能直接验证 (6.12) 式, 这可以用 $\theta(X)$ 和 $\theta(Y)$ 的局部描述来进行.

作为命题 6.4 的一个更为重要的应用, 我们考虑 $\theta(X)\circ d+d\circ\theta(X)$, 其中的 X 为光滑向量场. 因为 $d:\wedge^k\longrightarrow\wedge^{k+1}$, $\theta(X):\wedge^k\longrightarrow\wedge^{k-1}$, 得到 $\theta(X)\circ d+d\circ\theta(X):\wedge^k\longrightarrow\wedge^k$. 现在我们断言本节的主要公式成立:

$$D_X=\theta(X)\circ d+d\circ\theta(X).\tag{6.13}$$

由于 (6.13) 式的两边都是导数, 所以只要验证 (6.13) 式对于函数和线性形式成立即可. 如果 $f\in\wedge^0(M)$, 则 $\theta(X)f=0$, 因而由 (6.9) 式知等式 (6.13) 变为

$$D_Xf=\prec X,df\succ,$$

我们知道这是对的. 下一步必须验证 (6.13) 式对 $\omega\in\wedge^1(M)$ 成立. 由 (6.5) 式, 只需局部地验证 (6.13) 式. 如果我们记 $\omega=a_1dx^1+\cdots+a_ndx^n$, 由线性性, 只要对每项 a_idx^i 去验证 (6.13) 式即可. 因为 (6.13) 式两端都是导数, 我们有

$$D_X(a_idx^i)=(D_Xa_i)dx^i+a_i(D_Xdx^i)$$

和

$$[\theta(X)d + d\theta(X)](a_i dx^i) = [\theta(X)d + d\theta(X)](a_i)dx^i + a_i[\theta(X)d + d\theta(X)]dx^i.$$

由于我们已经对函数验证了 (6.13) 式, 所以只须对 dx^i 进行验证. 现在

$$D_X dx^i = dD_X x^i$$

而

$$[\theta(X)d + d\theta(X)]dx^i = d\theta(X)dx^i = d \prec X, dx^i \succ = dD_X x^i.$$

(6.13) 式的证明完成.

在许多场合为了方便会把 θ 这个字母作为他用. 因此有时我们采用记号

$$X\rfloor\omega = \theta(X)\omega.$$

符号 $\rfloor$ 被称为内乘积(interior product). $X\rfloor\omega$ 是形式 ω 与向量场 X 的内乘积. 如果 $\omega \in \wedge^k$, 则 $X\rfloor\omega \in \wedge^{k-1}$. 等式 (6.13) 则可重写为

$$D_X\omega = X\rfloor d\omega + d(X\rfloor\omega) \tag{6.14}$$

让我们看一看在局部坐标下的某些特殊情形中 (6.14) 式说了些什么. 如果 $\omega = a_1 dx^1 + \cdots + a_n dx^n$, 且

$$X = X^1\frac{\partial}{\partial x^1} + \cdots + X^n\frac{\partial}{\partial x^n},$$

则

$$d\omega = \sum da_i \wedge dx^i.$$

因而

$$\begin{aligned} X\rfloor d\omega &= \sum[(X\rfloor da_i) \wedge dx^i - da_i \wedge (X\rfloor dx^i)] \\ &= \sum_{i,j}\left(X^j\frac{\partial a_i}{\partial x^j}dx^i - X^j\frac{\partial a_j}{\partial x^i}dx^i\right), \end{aligned}$$

而

$$X\rfloor\omega = \sum a_j X^j,$$

故

$$d(X\rfloor\omega)=\sum X^j\left(\frac{\partial a_j}{\partial x^i}\right)dx^i+\sum a_j\left(\frac{\partial X^j}{\partial x^i}\right)dx^i.$$

所以

$$D_X\omega=\sum_i\left(\sum_j X^j\frac{\partial a_i}{\partial x^j}+a_j\frac{\partial X^j}{\partial x^i}\right)dx^i,$$

它与第九章的等式 (7.12′) 一致.

作为第二个例子, 设 $\Omega=adx^1\wedge\cdots\wedge dx^n$, 其中 $n=\dim M$. 于是 $d\Omega=0$, 故 (6.14) 式简化为

$$D_X\Omega=d(X\rfloor\Omega).$$

如果 $X=\sum X^i\left(\frac{\partial}{\partial x^i}\right)$, 则

$$\begin{aligned}X\rfloor\Omega&=\sum X^i\left(\frac{\partial}{\partial x^i}\right)\rfloor\omega\\&=\sum aX^i\left(\frac{\partial}{\partial x^i}\right)\rfloor(dx^1\wedge\cdots\wedge dx^n)\\&=\sum(-1)^{i-1}aX^idx^1\wedge\cdots\wedge dx^{i-1}\wedge dx^{i+1}\wedge\cdots\wedge dx^n,\end{aligned}$$

它仅仅是在第 4 节末尾所介绍的那个公式. 于是

$$D_X\Omega=d(X\rfloor\Omega)=\left(\sum\frac{\partial aX^i}{\partial x^i}\right)dx^1\wedge\cdots\wedge dx^n.$$

由于我们总能局部地把一个密度等同于一个 n - 形式, 即将 ρ 在 (U,α) 上等同于 $\rho_\alpha dx^1\wedge\cdots\wedge dx^n$, 我们便得到了第十章的命题 5.2 的另一个证明.

附录 I “向量分析”

在这里我们要列出本章所引进的概念与关于“向量分析”的书籍中出现的种种概念之间的关联, 尽管我们并没有机会去用它们.

在定向三维欧氏空间 $\mathbb{E}^3$ 中有许多等同关系, 我们可以用这一章中所引进的一些运算给它们一种特殊的形式.

首先, 像任意黎曼空间那样, 我们可以在 $\mathbb{E}^3$ 中把向量场等同于线性微分形式. 那么对任意函数 f, 我们可以把 df 看成一个向量场. 照这样, 它便被称作

grad$f$. 于是, 在 $\mathbb{E}^3$ 中用直角坐标 x, y, z 表示有

$$\operatorname{grad} f = \prec \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \succ,$$

其中, 我们也已经把在 $\mathbb{E}^3$ 上的向量场等同于 $\mathbb{E}^3$ - 值的函数.

其次, 由于 $\mathbb{E}^3$ 是定向的, 我们可以通过 $*$ - 算子 (作用于每个 T_x^*) 将 $\wedge^2(E^3)$ 与 $\wedge^1(E^3)$ 等同. 回忆 $*$, 它是由

$$*(dx \wedge dy) = dz, *(dx \wedge dz) = -dy, *(dy \wedge dz) = dx \tag{I.1}$$

给出.

特别地, 如果 $\omega_1 = \prec P, Q, R \succ = Pdx + Qdy + Rdz, \omega_2 = \prec L, M, N \succ = Ldx + Mdy + Ndz$, 我们便能引进所谓的 ω_1 与 ω_2 的 "向量乘积" . 它的定义是

$$\omega_1 \times \omega_2 = *(\omega_1 \wedge \omega_2),$$

并且 (用 (I,1)) 由

$$\prec P, Q, R \succ \times \prec L, M, N \succ = \prec QN - RM, RL - PN, PM - QL \succ$$

给出.

我们又引进算子

$$\operatorname{curl} \omega = *d\omega.$$

于是, 如果 $\omega = \prec P, Q, R \succ$, 我们便有

$$\operatorname{curl} \omega = \prec \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial R}{\partial y} \succ.$$

考虑 $\mathbb{E}^3$ 中的一个定向曲面, 即设 $\varphi: S \longrightarrow \mathbb{E}^3$. 设 Ω 为 S 上的体积形式, 它是由 φ 诱导的黎曼度量相关联的体积形式. 由定义, 如果 $\xi_1, \xi_2 \in T_x(S)$, 则

$$\Omega(\xi_1, \xi_2) = dV \prec \varphi_*\xi_1, \varphi_*\xi_2, n \succ$$

其中 dV 是 $\mathbb{E}^3$ 的体积元素而 n 是单位法向量. 这个等式的另一种写法是说

$$\Omega(\xi_1, \xi_2) = \sigma(\varphi_*\xi_1, \varphi_*\xi_2),$$

其中当我们把 n 看作微分形式时, $\sigma = *n$. 现设 $\overline{\omega}$ 为 $\mathbb{E}^3$ 中的形式, 并假设 $\varphi * \overline{\omega} = f\Omega$, f 为某个函数. 于是

$$f(x) = (\overline{\omega}, *n)(\varphi(x)).$$

因此

$$\int_S \varphi^*(\overline{\omega}) = \int_S f\Omega = \int_S (\overline{\omega}, *n)\Omega = \int (*\overline{\omega}, n)\Omega.$$

将此用于 $\overline{\omega} = d\omega$, 其中 $\omega = Pdx + Qdy + Rdz$, 我们则可重写斯托克斯定理为

$$\int_C \omega = \int_C Pdx + Qdy + Rdz = \int_S (\mathrm{curl}\ \omega, n)\Omega,$$

其中 S 为闭曲线 C 张成的某个曲面.

如果把这个评注用于 $\overline{\omega} = *\omega$ 及 $S = \partial D$, 由于 (对 $n = 3$) $** =\mathrm{id}$, 我们得到

$$\int (\omega,n)\Omega = \int_D d*\omega.$$

注意有

$$d*\omega = \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) dx \wedge dy \wedge dz,$$

这是我们写为 $\mathrm{div}\omega$ 的东西, 即

$$\mathrm{div}\omega = d*\omega.$$

(实际上它是 $\mathrm{div}\prec \omega, dV \succ$, 其中 dV 为体积元, 而我们将 ω 看为一个向量场.) 因此我们再次得到了散度定理. 注意到由于 d^2=0, 我们有

$$\mathrm{curl}(\mathrm{grad} f) = *ddf = 0$$

和

$$\mathrm{div}(\mathrm{curl}\ \omega) = d**d\omega = d^2\omega = 0.$$

附录 II　$\mathbb{E}^3$ 中曲面的初等微分几何

为了计算方便, 我们引进向量值微分形式的概念. 设 E 为向量空间, M 为微分流形. 我们所说的一个 p 次的 E - 值外微分形式 Ω 是一个规则, 它对每个 $x \in M$ 指定一个元 Ω_x, 而 Ω_x 是在 $T_x(M)$ 上的一个反称的 p 次 E - 值多重线性函数. 例如, 如果 $p = 0$, 则 E - 值零 - 形式正好是 M 上取值于 E 的一个函数. 一个 E - 值 1 - 形式是一个规则, 它对 M 的任意点的每个切向量 ξ 指定了一个 E 的元, 等等.

假设 E 是有限维的, $\{e_1,\cdots,e_N\}$ 为 E 的一组基. 设 $\Omega_1,\cdots,\Omega_N$ 为 (实值) p - 形式. 我们于是可以考虑 E - 值 p - 形式 $\Omega=\Omega_1e_1+\cdots+\Omega_Ne_N$, 其中对任意 p 个 $T_x(M)$ 中的向量 $\xi_1,\cdots,\xi_p$, 我们有

$$\Omega_x(\xi_1,\cdots,\xi_p)=\Omega_{1x}(\xi_1,\cdots,\xi_p)e_1+\cdots+\Omega_{Nx}(\xi_1,\cdots,\xi_p)e_N.$$

反之, 如果 Ω 是 E - 值形式, 则实形式 $\Omega_1,\cdots,\Omega_N$ 可以由上面的等式定义. 简而言之, 一旦 N 维向量空间的基被选定, 则给出一个 E - 值微分形式就等同于给出 N 个实形式, 并且我们可以写为

$$\Omega=\sum_1^N\Omega_ie_i\qquad\text{或}\qquad\Omega=\prec\Omega_1,\cdots,\Omega_N\succ.$$

对 E - 值形式的局部描述的规则以及转移的法则与实形式的相似, 我们不再赘述. 为简明起见, 我们只限于 E 为有限维的情形, 尽管这个假定条件在大部分场合是不必要的.

如果 ω 是 p 次实微分形式, 我们可以以显见的方式定义 $\omega\wedge\Omega$. 利用一组基时, 如果 $\Omega=\prec\Omega_1,\cdots,\Omega_N\succ$, 则 $\omega\wedge\Omega=\prec\omega\wedge\Omega_1,\cdots,\omega\wedge\Omega_N\succ$.

更一般地, 设 E 和 F 为 (有限维的) 向量空间, 并设 $\#:E\times F\longrightarrow G$ 为一个双线性映射, 其中的 G 是第三个向量空间. 又设

$$\{e_1,\cdots,e_N\}$$

是 E 的基, $\{f_1,\cdots,f_M\}$ 为 F 的基, $\{g_1,\cdots,g_L\}$ 为 G 的基. 假设映射 $\#$ 由

$$\#\{e_i,f_i\}=\sum a_{ij}^kg_k$$

给出. 于是, 如果 $\omega=\sum\omega_ie_i$ 是一个 E - 值形式, $\Omega=\sum\Omega_jf_j$ 是个 F - 值形式, 我们定义 G - 值形式 $\omega\wedge\Omega$ 为

$$\omega\wedge\Omega=\sum_k\left(\sum_{i,j}a_{ij}^k\omega_i\wedge\Omega_j\right)g_k.$$

容易验证此定义不依赖于特定基的选取.

我们将要把这个概念主要地用于两种情形. 首先, 我们对于 $E=F$ 和 $G=\mathbb{R}$ 的情形有兴趣, 因而 $\#$ 是 E 上的双线性形式. 假设 $\#$ 是个内积而 $e_1,\cdots,e_N$ 为法正交基. 于是我们将其写成 $(\omega\wedge\Omega)$ 以提醒我们用的是内积. 如果

$$\omega=\prec\omega_1,\cdots,\omega_N\succ,\qquad\Omega=\prec\Omega_1,\cdots,\Omega_N\succ;$$

则

$$(\omega \wedge \Omega) = \sum \omega_i \wedge \Omega_i.$$

注意在这种情形中, 如果 ω 为 p - 形式, Ω 为 q - 形式, 则

$$(\omega \wedge \Omega) = (-1)^{pq}(\Omega \wedge \omega),$$

这与实形式情形一样.

我们感兴趣的第二种情形是 $F = G$ 而 $E = \mathrm{Hom}(F)$, # 是赋值映射: 线性变换在 F 的一个向量上的取值给出了 F 的另一个值. 这一回, 选取 F 的一组基便决定了 $\mathrm{Hom}(F)$ 的一组基, 故可将 ω 看成是实值微分形式的一个矩阵. 如果 $\omega = (\omega_{ij}), \Omega = \prec \Omega_1, \cdots, \Omega_M \succ$, 则

$$\omega \wedge \Omega = \prec \sum \omega_{1j} \wedge \Omega_j, \cdots, \sum \omega_{Mj} \wedge \Omega_j \succ.$$

算子 d 对于向量值形式的意思与实形式情形相同, 满足同一规则. 因此, 如果 $\Omega = \prec \Omega_1, \cdots, \Omega_N \succ$, 则 $d\Omega = \prec d\Omega_1, \cdots, d\Omega_N \succ$, 且

$$d(\omega \wedge \Omega) = (d\omega \wedge \Omega) + (-1)^p(\omega \wedge d\Omega),$$

其中 ω 为 p 次 E - 值形式, Ω 为 F - 值形式.

我们要用向量值形式的概念去阐述 (大部分以习题形式) 关于 $\mathbb{E}^3$ 中定向曲面的一些几何的初等内容. 设 M 为定向二维流形, φ 为 M 到 $\mathbb{E}^3$ 中的可微映射. 假设 φ_* 在 M 的任何点都不是奇异的, 即 φ 是个浸入. 因此在每点 $p \in M$, 空间 $\varphi_*(T_p(M))$ 是 $T_{\varphi(p)}(\mathbb{E}^3)$ 的一个二维子空间. 由于我们可以把 $T_{\varphi(p)}(\mathbb{E}^3)$ 等同于 $\mathbb{E}^3$, 所以可以把 $\varphi_*(T_p(M))$ 看为 $\mathbb{E}^3$ 的二维子空间 (见图 11.15). 因为 M 是

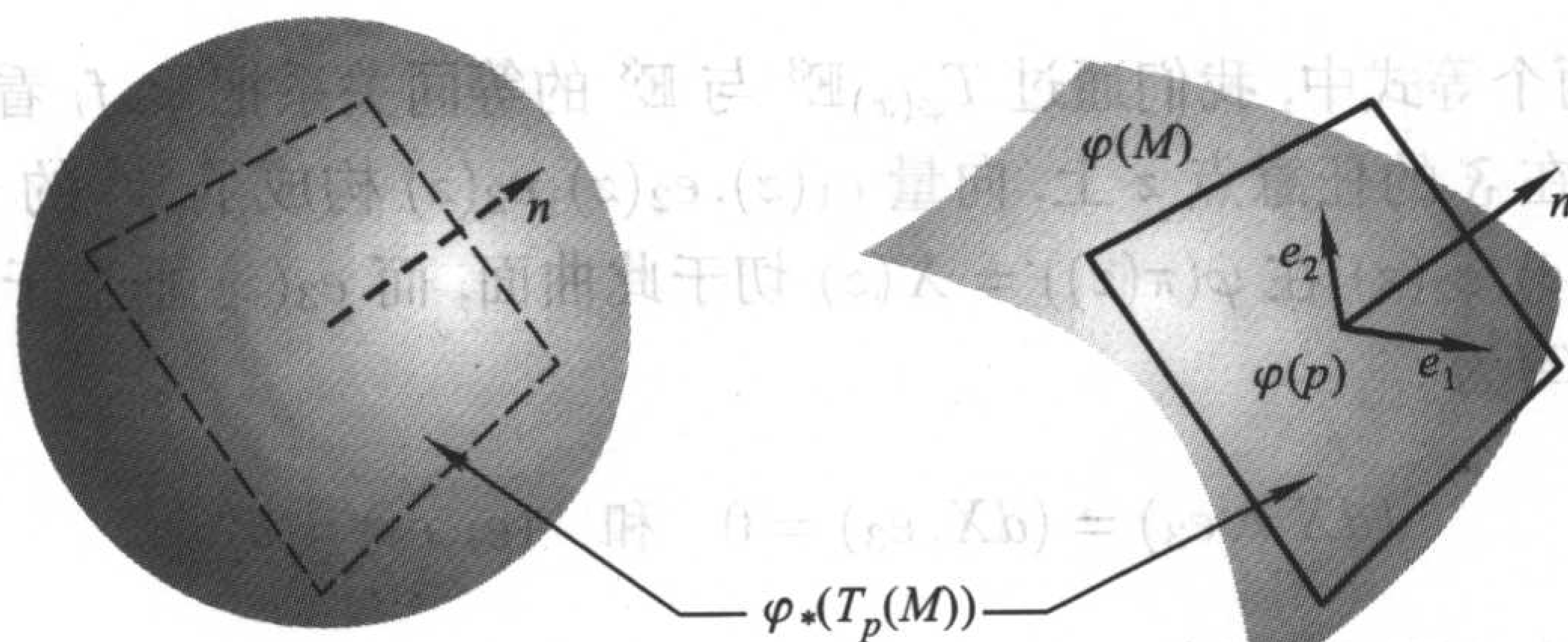


图 11.15

定向的, 从而切平面 $\varphi_*(T_p(M))$ 也是定向的. 于是有一个惟一的单位向量垂直于切平面, 它连同切平面的定向基给出了 $\mathbb{E}^3$ 的一个定向基. 称这个向量为法向量, 记为 $n(p)$. 我们可以把 n 考虑为 M 上一个 $\mathbb{E}^3$ - 值的函数. 由于 $\| n \|= 1$, 从而能把 n 看成是 M 到单位球面的映射. 注意 $\varphi(M)$, 它位于 $\mathbb{E}^3$ 中一个固定平面中的充要条件是 $n=$ 常值 ($n=$ 此平面的法向量). 因此我们可以预料到 n 的变化对于描述曲面 $\varphi(M)$ 是如何“弯曲”方面是有用的.

设 Ω 为 M 上对应于由 φ 诱导的黎曼度量的 (定向) 面积形式. 设 Ω_S 为单位球面上的 (定向) 面积形式. 于是 $n^*(\Omega_S)$ 是 M 上的 2 - 形式, 因此可以记为

$$n^*\Omega_S = K\Omega.$$

函数 K 被称作曲面 $\varphi(M)$ 的高斯曲率. 注意, 当 $\varphi(M)$ 在一个平面中时 $K = 0$. 另外如果 $\varphi(M)$ 是圆柱面也有 $K = 0$(见习题).

对任意一个具黎曼度量的定向二维流形, 我们让 $\mathfrak{F}$ 代表 M 的所有切空间的所有定向基的集合. 因此 $\mathfrak{F}$ 的一个元素由 $\prec f_1, f_2 \succ$ 给出, 其中 $\prec f_1, f_2 \succ$ 是对某个 $x \in M$ 的 $T_x(M)$ 的法正交基. 注意, 因为定向及 $f_2 \perp f_1$, 故 f_2 被 f_1 所决定. 于是我们可以把 $\mathfrak{F}$ 考虑为所有具单位长度的切向量的空间. 对每个 $x \in M$, 所有单位向量的集合正好是一个圆. 我们留给读者去验证 $\mathfrak{F}$ 事实上是一个三维流形. 对每个 $\prec f_1, f_2 \succ$ 派定给点 x 的映射记为 π, 这里的 $\prec f_1, f_2 \succ$ 是在 x 的一组法正交基. 读者还应该验证 π 是 $\mathfrak{F}$ 到 M 上的一个可微映射.

在当前的情形中, 即度量来自浸入 φ 时, 我们定义在 $\mathfrak{F}$ 上的数个向量值函数 X, e_1, e_2, e_3 如下:

$$X = \varphi \circ \pi,$$
$$e_1(\prec f_1, f_2 \succ) = \varphi_* f_1, e_2(\prec f_1, f_2 \succ) = \varphi_* f_2, e_3 = n \circ \pi.$$

(在中间的两个等式中, 我们通过 $T_{\varphi(x)}\mathbb{E}^3$ 与 $\mathbb{E}^3$ 的等同关系把 $\varphi_* f_i$ 看作 $\mathbb{E}^3$ 中的元素.) 因此在 $\mathfrak{F}$ 的任意点 z 上, 向量 $e_1(z), e_2(z), e_3(z)$ 构成了 $\mathbb{E}^3$ 的一组法正交基, 其中 $e_1(z), e_2(z)$ 在 $\varphi(\pi(z)) = X(z)$ 切于此曲面, 而 $e_3(z)$ 垂直于此曲面. 我们因此可以写为

$$(dX \wedge e_3) = (dX, e_3) = 0 \quad 和 \quad (e_i, e_j) = \delta_i^j.$$

由第一个方程, 可写出

$$dX = \omega_1 e_1 + \omega_2 e_2, \tag{II.1}$$

其中

$$\omega_1 = (dX, e_1), \quad \omega_2 = (dX, e_2)$$

为在 $\mathfrak{F}$ 上定义的 (实值) 线性微分形式.

相似地, 我们定义形式 ω_{ij} 为

$$\omega_{ij} = (de_i, e_j).$$

将 d 用到算式 $(e_i, e_j) = \delta_{ij}$ 上表明

$$\omega_{ij} = -\omega_{ji}. \tag{II.2}$$

如果将 d 用于 (II.1), 我们得到

$$0 = ddX = d\omega_1 e_1 - \omega_1 \wedge de_1 + d\omega_2 e_2 - \omega_2 \wedge de_2,$$

分别取此方程与 e_1 和 e_2 的内积, 表明了 (因为 $\omega_{11} = 0, \omega_{22} = 0$)

$$d\omega_1 = \omega_2 \wedge \omega_{21}, \quad d\omega_2 = \omega_1 \wedge \omega_{12}. \tag{II.3}$$

如果把 d 作用于等式

$$de_i = \sum_j \omega_{ij} e_j,$$

我们得到

$$0 = \sum (d\omega_{ij} e_j - \omega_{ij} \wedge de_j),$$

如果再取与 e_j 的内积, 我们得到了

$$d\omega_{ij} = \sum_k \omega_{ik} \wedge \omega_{kj}. \tag{II.4}$$

如果又将 d 施与方程 $(dX, e_3) = 0$, 得出了

$$0 = d(dX, e_3) = (dX, de_3) = (\omega_1 e_1 + \omega_2 e_3, \omega_{31} e_1 + \omega_{32} e_2),$$

它表示

$$\omega_1 \wedge \omega_{31} + \omega_2 \wedge \omega_{32} = 0. \tag{II.5}$$

我们现在要解释这些等式. 设 $Z = \prec f_1, f_2 \succ$ 为 $\mathfrak{F}$ 中一点. 对任意 $\xi \in T_x(\mathfrak{F})$ 我们有

$$\langle \xi, dX \rangle = \langle \xi, d\pi^* \varphi \rangle = \langle \xi, \pi^* d\varphi \rangle = \langle \pi_* \xi, d\varphi \rangle = \varphi_* (\pi_* \xi).$$

因此, 由于这个度量的定义使得 φ_* 为等距, 故有

$$\begin{aligned}\langle\xi,\omega_1\rangle &= (\varphi_*\pi_*\xi,e_1) = (\varphi_*\pi_*\xi,\varphi_* f_1)\\ &= (\pi_*\xi, f_1).\end{aligned}\tag{II.6}$$

换句话说, $\langle\xi,\omega_1\rangle$ 与 $\langle\xi,\omega_2\rangle$ 是 $\pi_*\xi$ 关于基 $\prec f_1, f_2 \succ$ 的分量. 如果 η 是在 z 的另一个切向量, 那么 $\omega_1\wedge\omega_2(\xi,\eta)$ 便是由 $\pi_*\xi$ 和 $\pi_*\eta$ 张成的平行四边形的 (定向) 面积. 换句话说

$$\omega_1\wedge\omega_2 = \pi^*\Omega,\tag{II.7}$$

其中的 Ω 是 M 上的定向面积.

相似地,

$$\langle\xi,de_3\rangle = n_*\pi_*\xi,\tag{II.8}$$

从而我们有

$$\begin{aligned}n_*\pi_*\xi &= \langle\xi,\omega_{31}\rangle e_1 + \langle\xi,\omega_{32}\rangle e_2\\ &= \langle\xi,\omega_{31}\rangle\varphi_* f_1 + \langle\xi,\omega_{32}\rangle\varphi_* f_2.\end{aligned}\tag{II.9}$$

由于我们可以把 e_1 和 e_2 看为单位球面的切空间的法正交基, 我们的结论是说, $\omega_{31}\wedge\omega_{32}(\xi,\eta)$ 是由 $n_*\pi_*\xi, n_*\pi_*\eta$ 在单位球面上张成的平行四边形的定向面积. 因此

$$\omega_{31}\wedge\omega_{32} = \pi^* n^*\Omega_S = \pi^* K\Omega = K\omega_1\wedge\omega_2.$$

设

$$\begin{bmatrix} a & b\\ b^1 & c\end{bmatrix}$$

是线性变换 $n_*: T_x(M)\longrightarrow T_{n(x)}(S^2)$ 在 $T_x(M)$ 的基 $\prec f_1, f_2\succ$ 和 $T_{n(x)}(S^2)$ 的基 $\prec e_1, e_2\succ$ 下的矩阵. 那么比较 (II.6) 与 (II.9) 表明

$$\omega_{31} = a\omega_1 + b^1\omega_2, \omega_{32} = b\omega_1 + c\omega_2.\tag{II.10}$$

如果将此代入 (II.5) 则得到 $b = b^1$, 即 n_* 的矩阵是对称的. 这暗示了它对应于某种几何意义的对称双线性形式. 换句话说, 我们需要考虑二次型

$$a\omega_1^2 + 2b\omega_1\omega_2 + c\omega_2^2$$

[在这里可以理解为这是一个 $T_z(\mathfrak{F})$ 上的二次型, 它对任意 $\xi\in T_z(\mathfrak{F})$ 指定了数 $a\langle\xi,\omega_1\rangle^2 + 2b\langle\xi,\omega_1\rangle\langle\xi,\omega_2\rangle + c\langle\xi,\omega_2\rangle^2$].

习题

II.1 证明

$$a\langle\xi,\omega_1\rangle^2+2b\langle\xi,\omega_1\rangle\langle\xi,\omega_2\rangle+c\langle\xi,\omega_2\rangle^2=\langle\varphi_*\pi_*\xi,n_*\pi_*\xi\rangle$$

II.2 对每个$\zeta\in T_x(M)$ 指定数 $(\varphi_*\zeta,n_*\zeta)$ 的二次型被称作此曲面的第二基本形式. 我们记其为 $\mathrm{II}(\zeta)$(通常称之为第一基本形式的在我的术语下就是 $\|\zeta\|^2$.) 设 C 是条光滑曲线, 满足 $C'(0)=\zeta$. 证明

$$\mathrm{II}(\zeta)=-\left(\frac{d^2\varphi\circ C}{dt^2}(0),n(x)\right).$$

因此 $\mathrm{II}(\zeta)$ 度量了曲线 $\varphi\circ C$ 在 n 方向弯曲了多少. 假设我们选取 C 使 $\varphi\circ C$ 在由 $\varphi_*\xi$ 与 $n(x)$ 张成的平面中.[几何地看, 这等于考虑由 $\varphi_*\xi$ 和 $n(x)$ 张成的平面与曲面相交得到的此曲面上的曲线.] 证明 $\mathrm{II}(\zeta)$ 是此平面曲线的曲率.

在这个意义下, 第二基本形式告诉我们此曲面在 ζ 方向上的弯曲有多少.

注意有

$$K=\det\begin{bmatrix}a&b\\b&c\end{bmatrix}.$$

设 λ_1 和 λ_2 为此矩阵 $\begin{bmatrix}a&b\\b&c\end{bmatrix}$ 的特征值. 因此对于 $\|\zeta\|=1$ 有

$$\lambda_1=\max\mathrm{II}(\zeta),\qquad\lambda_2=\min\mathrm{II}(\zeta).$$

如果 $\lambda_1\neq\lambda_2$, 则有两个正交的特征向量, 称其为此曲面的主曲率方向.(注意, 它们必定正交, 原因在于它们是对称矩阵的特征向量.)

如果 A 是 $\mathbb{E}^3$ 的一个欧几里得运动, 则 $\psi=A\circ\varphi$ 是 M 的另一个浸入, 而且容易验证由 ψ 诱导的黎曼度量和与 ψ 关联的第二基本形式都与 φ 相联系的那些是相互一致的. 但其逆却不是那么显而易见的: *如果 ψ 和 φ 诱导出相同的度量和相同的第二基本形式, 则有某个欧几里得运动 A 使得 $\psi=A\circ\varphi$.* 我们不打算证明这个结果, 尽管它是我们已经建立了的那些东西的十分容易的推论.

我们已经看到以几何术语表达的 $\omega_1,\omega_2,\omega_{31},\omega_{32}$ 的意思. 让我们现在来解释剩下的一个形式 ω_{12}.

设 γ 是 M 上的可微曲线. 沿着 γ 的一个单位向量的可微族 $f_1(\cdot)$[其中 $f_1(s)\in T_{\gamma(s)}(M)$] 是与 $\mathfrak{F}$ 中的一条满足 $\pi\circ C=\gamma$ 的一条曲线 C 意思相同. [即

$C(s) = \prec f_1(s), f_2(s) \succ$.] 我们称族 $f_1(s)$ 是平行的是说这些单位向量全都按曲面在三维空间中法向量那样变化. 换句话说, $f_1(s)$ 是平行的则表示向量

$$\frac{d_{\varphi_*\gamma(s)}(f_1(s))}{ds}$$

对所有的 s 垂直于 $\varphi(M)$. 我们来看看该如何表达这个条件. 设 ξ_s 为曲线 C 在 $C(s)$ 的切向量. 于是由 de_1 的定义有

$$\frac{d}{ds}\varphi_{*\gamma(s)}(f_1(s)) = \langle \xi_s, de_1 \rangle.$$

注意有 $(\langle \xi_s, de_1 \rangle, e_1(C(s))) = 0, (\langle \xi_s, de_1 \rangle, e_2(C(s))) = \langle \xi_s, \omega_{12} \rangle$. 现在 e_1 和 e_2 张成了 $\varphi(M)$ 的切空间, 所以说 $f_1(\cdot)$ 平行等于说 $\langle \xi_s, \omega_{12} \rangle = 0$.

因此 $f_1(s)$ 沿 γ 平行当且仅当 $\langle \xi_s, \omega_{12} \rangle = 0$.

设 M 和 $\overline{M}$ 为具黎曼度量的二维流形. 设 $u : M \longrightarrow \overline{M}$ 为等距的可微映射. 设 $\mathfrak{F}$ 为 M 的法正交基的流形, $\overline{\mathfrak{F}}$ 为 $\overline{M}$ 的法正交基的流形. 于是 u 诱导出映射 $\widetilde{u} : \mathfrak{F} \longrightarrow \overline{\mathfrak{F}}$ 为

$$\widetilde{u}(\prec f_1, f_2 \succ) = \prec u_* f_1, u_* f_2 \succ .$$

设 ω_1 为 $\mathfrak{F}$ 上的微分形式, 它如同 (II.6) 那样, 由

$$\langle \xi, \omega_1 \rangle = \langle \pi_* \xi, f_1 \rangle$$

给出, 其中 $\xi \in T_z(\mathfrak{F})$, 而 $z = \prec f_1, f_2 \succ$, 以及相应的关于 $\omega_2, \widetilde{\omega}_1$ 和 $\widetilde{\omega}_2$ 的定义. 于是对任意 $\xi \in T_z(\mathfrak{F})$ 我们有

$$\begin{aligned}\langle \xi, \widetilde{u}_* \widetilde{\omega}_1 \rangle &= \langle \widetilde{u}_* \xi, \omega_1 \rangle = (\widetilde{\pi}_* \widetilde{u}_* \xi, u_* f_1) = (u_* \pi_* \xi, u_* f_1) \\ &= (\pi_* \xi, f_1) = \langle \xi, \omega_1 \rangle,\end{aligned}$$

其中用到了 $\widetilde{\pi} \circ \widetilde{u} = u \circ \pi$. 换句话说,

$$\widetilde{u}^* \widetilde{\omega}_1 = \omega_1, \qquad \widetilde{u}^* \widetilde{\omega}_2 = \omega_2.$$

现在假定有来自浸入 φ 和 $\widetilde{\varphi}$ 的 M 和 $\overline{M}$ 上的度量. 于是得到了形式 ω_{ij} 和 $\widetilde{\omega}_{ij}$. 由 (II.3) 我们得到

$$\widetilde{u}^*(\widetilde{\omega}_1 \wedge \widetilde{\omega}_{21}) = \widetilde{u}^* d\widetilde{\omega}_1 = d(\widetilde{u}^* \widetilde{\omega}_1) = d\omega_1 = \omega_1 \wedge \omega_{21}.$$

因此

$$\omega_1 \wedge \widetilde{u}^* \widetilde{\omega}_{21} = \omega_1 \wedge \omega_{21} \qquad \text{和} \qquad \omega_2 \wedge \widetilde{u}^* \widetilde{\omega}_{12} = \omega_2 \wedge \omega_{12},$$

或者

$$(\widetilde{u}^*\widetilde{\omega}_{12}-\omega_{12})\wedge\omega_1=0 \qquad 和 \qquad (\widetilde{u}^*\omega_{12}-\omega_{12})\wedge\omega_2=0.$$

因为微分形式 ω_1 和 ω_2 是线性无关的, 这只有当

$$\widetilde{u}^*\widetilde{\omega}_{12}=\omega_{12}$$

才能成立.

换句话说, 如果两个曲面 $\varphi(M)$ 和 $\overline{\varphi}(M)$ 为等距的, 我们则有"相同"的 ω_{12}, 即有相同的"平行向量场"的概念. 注意到圆柱的一片与平面的一片是等距, 虽然说它们在欧几里得运动下不是全等的. 用不同的话来表达, 可以说成是在形式 ω_{13} 和 ω_{23} 依赖于此曲面是怎样浸入 $\mathbb{E}^3$ 的同时, 形式 ω_{12} 却只依赖于此浸入所诱导的黎曼度量.

现在我们有 (II.4):

$$d\omega_{12}=\omega_{13}\wedge\omega_{23}=-\pi^*\Omega_s=-K\Omega.$$

由此我们得出结论说, 高斯曲率 K 不依赖于浸入而只依赖于此浸入诱导的黎曼度量.

因为 ω_{12} 不依赖于 φ, 我们应该能够对任意具有黎曼度量的二维流形定义它. 注意前面的论证, 它指出 ω_{12} 惟一地被等式 (II.4) 决定. 因此只要在一个坐标邻域中构造一个满足 (11.4) 的 ω_{12} 即可. 然后由惟一性推出任何两个这样的东西重合, 从而给出了确定的形式. 设 U 是 M 的一个坐标邻域, $\psi: U\longrightarrow\mathfrak{F}$ 为可微映射使得 $\pi\circ\psi=\mathrm{id}$. 于是 ψ 对每个 $x\in U$ 以可微的方式派定一个基 $\prec f_1,f_2\succ$.(构造 ψ 的一种可能的方法是对向量场 $\prec\dfrac{\partial}{\partial x^1},\dfrac{\partial}{\partial x^2}\succ$ 使用法正交化步骤.)

一旦我们选定了 ψ, 那么 T_x 的任一组基与 $\psi(x)$ 只相差一个旋转. 如果我们让 τ 表示给出这个旋转的 (角) 坐标 (故 τ 只在 mod2π 下定义), 我们则可以使用 U 上的局部坐标连同 τ 一起作为 $\pi^{-1}(U)$ 的坐标. 更准确地, 如果 x^1,x^2 为 U 上的局部坐标, 我们定义 y^1,y^2,τ 为

$$y^1=x^1\circ\pi, \qquad y^2=x^2\circ\pi,$$

而对 $z=\prec e_1,e_2\succ,\tau(z)$ 由

$$e_1=\cos\tau(z)f_1+\sin\tau(z)f_2, \qquad e_2=-\sin\tau(z)f_1+\cos\tau(z)f_2, \tag{II.11}$$

其中 $\prec f_1, f_2 \succ = \psi(x), \prec e_1, e_2 \succ \in T_x(M)$.

现在设

$$\theta_1 = \psi^*(\omega_1), \qquad \theta_2 = \psi^*(\omega_2),$$

于是 θ_1 和 θ_2 是定义于 U 上的形式, 事实上在每个 $x \in M$ 上是 $\psi(x)$ 的对偶基. 如果令

$$\alpha_1 = \pi^*\theta_1, \qquad \alpha_2 = \pi^*\theta_2,$$

则 (II.11) 给出了

$$\omega_1 = \cos\tau\alpha_1 + \sin\tau\alpha_2, \qquad \omega_2 = -\sin\tau\alpha_1 + \cos\tau\alpha_2.$$

注意到有

$$\omega_1 \wedge \omega_2 = \alpha_1 \wedge \alpha_2.$$

定义 M 上的函数 l_1, l_2 为

$$d\theta_1 = l_1\theta_1 \wedge \theta_2, \qquad d\theta_2 = l_2\theta_1 \wedge \theta_2.$$

设 $k_1 = l_1 \circ \pi, k_2 = l_2 \circ \pi$, 故

$$d\alpha_1 = k_1\alpha_1 \wedge \alpha_2, \qquad d\alpha_2 = k_2\alpha_1 \wedge \alpha_2$$

现在有

$$\begin{aligned} d\omega_1 &= -\sin\tau d\tau \wedge \alpha_1 + \cos\tau d\tau \wedge \alpha_2 + (k_1\cos\tau + k_2\sin\tau)\alpha_1 \wedge \alpha_2, \\ d\omega_2 &= -\cos\tau d\tau \wedge \alpha_1 - \sin\tau d\tau \wedge \alpha_2 + (k_2\cos\tau - k_1\sin\tau)\alpha_1 \wedge \alpha_2. \end{aligned}$$

因为 $\omega_1 \wedge \omega_2 = \alpha_1 \wedge \alpha_2$, 我们可重写这些等式为

$$\begin{aligned} d\omega_1 &= (d\tau + (k_1\cos\tau + k_2\sin\tau)\omega_1) \wedge \omega_2, \\ d\omega_2 &= -(d\tau - (k_2\cos\tau - k_1\sin\tau)\omega_2) \wedge \omega_1. \end{aligned}$$

因此我们看出形式

$$\begin{aligned} \omega_{12} &= d\tau + (k_1\cos\tau + k_2\sin\tau)\omega_1 + (-k_1\sin\tau + k_2\cos\tau)\omega_2 \\ &= d\tau + k_1\alpha_1 + k_2\alpha_2 \end{aligned}$$

满足所要的等式.

像前面那样, 在任意二维黎曼流形上我称一族单位向量沿曲线 γ 平行是说 $\langle \xi_s, \omega_{12} \rangle = 0$. 用此平行移动的定义我们可以叙述下面的定理了.

定理 II.1　设 γ 为 M 上的任意一条可微曲线. 给出单位向量 $g_1 \in T_{\gamma(0)}(M)$, 则存在沿 γ 的惟一的平行单位向量族 $g_1(s)$, 满足 $g_1(0) = g_1$. 如果 $g_1'(0)$ 是 $T_{\gamma(0)}(M)$ 中一个不同于 g_1 的另一个单位向量, 并与 g_1 相差一个 σ 角, 则对所有 $s, g'(s)$ 与 $g_1(s)$ 也相差同样的角 σ.

证明　显然只要对完全位于一个坐标分图表中的曲线 γ 去证明定理即可 (如果必要可将 γ 分割成若干小段). 于是我们可以利用 ω_{12} 的局部表示.

重写沿 $\gamma(s)$ 的平行移动的条件. 在局部坐标下, 单位向量 $g_1(s)$ 由函数 $\tau(s)$ 给出,

$$g_1(s) = \cos\tau(s) f_1(\gamma(s)) - \sin\tau(s) f_2(\gamma(s)).$$

于是

$$\begin{aligned}\langle \xi_s, \omega_{12}\rangle &= \langle \xi_s, d\tau\rangle + \langle \xi_s, k_1\alpha_1 + k_2\alpha_2\rangle \\ &= \frac{d\tau(s)}{ds} + \langle \xi_s, \pi^*(k_1\theta_1 + k_2\theta_2)\rangle \\ &= \frac{d\tau(s)}{ds} - \langle \pi_*\xi_s, k_1\theta_1 + k_2\theta_2\rangle.\end{aligned}$$

但是 $\pi_*\xi_s = \zeta_s$ 是 γ 在 $\gamma(s)$ 的切向量. 从而

$$\langle \xi_s, \omega_{12}\rangle = \frac{d\tau(s)}{ds} - F_\gamma(s),$$

其中 $F_\gamma(s) = \langle \zeta_s, k_1\theta_1 + k_2\theta_2\rangle$ 是只依赖于 s 的一个函数. 特别地, $g_1(s)$ 是平行的充要条件是

$$\frac{d\tau(s)}{ds} = F_\gamma(s).$$

由此我们知道, 给出 $g_1(0)$ 后, 存在一条从 $g_1(0)$ 出发的惟一平行族 $g_1(s)$. 另外, 如果 $g_1'(0)$ 是在 $\gamma(0)$ 的第二个单位向量, 则 $g_1(s)$ 与 $g_1'(s)$ 之间的角等于 $g_1(0)$ 与 $g_1'(0)$ 之间的角. 那么, 平行移动保持了角不变, 定理得证. □

注意, 如果 M 是 (局部等距于) 欧几里得空间, 则我们可以选取

$$f_1 = \frac{\partial}{\partial x^1}, \qquad f_2 = \frac{\partial}{\partial x^2},$$

从而 $\theta_1 = dx^1, \theta_2 = dx^2$. 这时 $k_1 = k_2 = 0$ 而 τ 正是 g_1 与 $\dfrac{\partial}{\partial x^1}$ 即 x_1 - 轴所构成的角. 因此 $\omega_{12} = -d\tau$. 于是平行移动的条件成了 $\dfrac{d\tau}{ds} = 0$, 它同于欧氏几何中通常的平行概念. 但是注意一下, 在欧氏空间中平行性不依赖于曲线 γ. 在一般情形这是不对的.

习题

11.3 设 γ_1, γ_2 为连接南北极的, S^2 上的两个大圆弧. 假定 γ_1 和 γ_2 在极点正交. 设 ζ 是在北极的一个切向量. 比较通过 γ_1 和 γ_2 到南极的移动.

设 M 为任意一个二维黎曼流形. 对 M 上任一条曲线 γ 有一个沿 γ 选择单位向量的明显方法: 让 $g_1(s)$ 就是 γ 在 $\gamma(s)$ 的单位切向量. 因此对 M 上每条曲线 γ 我们得到 $\mathfrak{F}$ 上一条曲线, 记其为 $\tilde{\gamma}$.[这时 $\tilde{\gamma} = (\gamma(s), g_1(s), g_2(s))$, $g_1(s)$ 是 $\gamma(s)$ 的切向量.]

称形式 $\tilde{\gamma}^*(\omega_{12})$ 为 γ 的**测地曲率**形式.[在欧几里得情形这恰好是通常的曲率(见习题).]

让我们考虑那些使测地曲率为零的曲线, 即那些切向量相互平行的曲线. 我们称这种曲线为关于所给黎曼度量的**测地线**. 注意, 在局部坐标下一条曲线为测地线是由一个二阶微分方程给出的. 因此测地线 $C(\cdot)$ 被 $C(t), C'(t)$ 在任意固定 t 的值所惟一确定. 在第十三章中我们要用"测地线"来表示一条局部使长度极小化的曲线. 下面的几个习题的目的是要证明在现在意义下的测地线具有这个性质.

习题

11.4 设 x, y 为 $U \subset M$ 上的局部坐标. 通过曲线 $y = 0$(即在局部坐标下的 x - 轴) 上每点, 构造正交于此曲线的惟一的测地线 (见图 11.16). 设 s 为沿此测地线的弧长参数, 从而通过 $(u, 0)$ 的测地线由

$$(y(u, s), x(u, s))$$

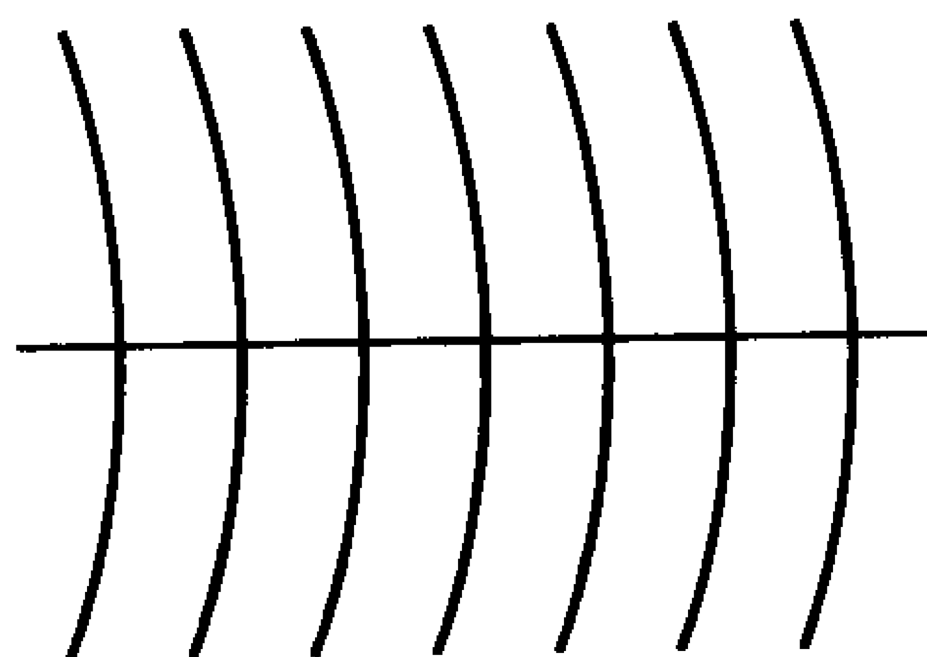


图 11.16

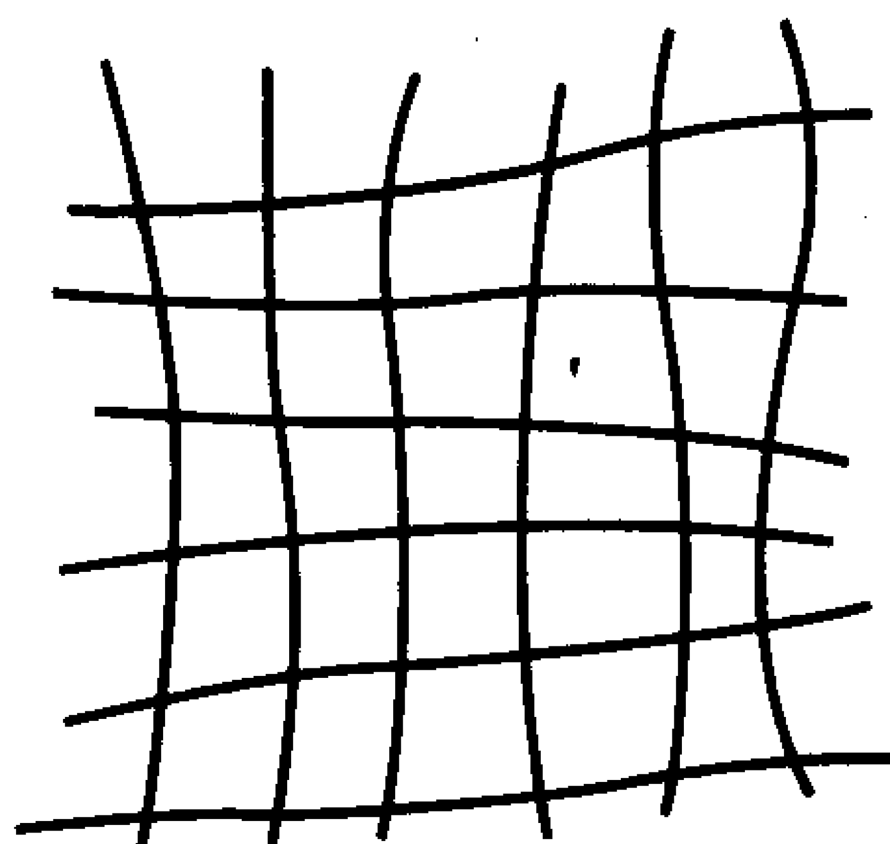


图 11.17

给出. 证明映射 $(u,s)\longmapsto(y(u,s),x(u,s))$ 在 (0,0) 有非零雅可比, 因而在某个开子集 $U'\subset U$ 上定义了一个坐标系.

II.5 我们要做进一步的坐标变换. 设 Y 为 U' 上的向量场, 其由下面的性质所定义:

$$\| Y \|=1, \left(Y,\frac{\partial}{\partial s}\right)=0, \langle Y,du\rangle>0.$$

因而 Y 正交于测地线 $u=$ 常数并且指向 u 的增大方向. 我们来考虑这个向量场的解曲线, 它是沿测地线 $u=0$ 上的初始位置作为参数的. 这就是说, 设 v 为沿测地线 $u=0$ 的弧长参数, 并考虑映射

$$(u,v)\longmapsto(u,s(u,v)),$$

中的 $s(u,v)$ 是 Y 经过 $(0,v)$ 的解曲线与由 u 给出的测地线的交点的 s - 坐标 (见图 11.17). 存在性定理, 对参数的光滑地依赖性连同曲线 $u=0$ 和 $s=0$ 已经是正交的这个事实, 保证了我们能够找到某个邻域 W 使得 (u,v) 为 W 上的坐标. 因此我们便构造了坐标系使得曲线 $u=$ 常数为测地线, 而曲线 $u=$ 常数和 $v=$ 常数正交. 这样一个坐标系被称为*测地平行坐标系*.

II.6 设 (u,v) 为 $U\subset M$ 上的坐标系, 其中 $\left(\frac{\partial}{\partial u},\frac{\partial}{\partial v}\right)\equiv 0$, 故此度量有形式

$$ds^2=pdu^2+qdv^2.$$

对 $\frac{\partial}{\partial u},\frac{\partial}{\partial v}$ 的法化定义了一个标架 ψ, 使得 $\psi(x)=\prec f_1,f_2\succ$, 其中 $f_1=\left(\frac{\partial}{\partial u}\right)\Big/\left\|\frac{\partial}{\partial u}\right\|$ 和 $f_2=\left(\frac{\partial}{\partial v}\right)\Big/\left\|\frac{\partial}{\partial v}\right\|$. 证明 $\theta_1=pdu,\theta_2=qdv$ 给出了形式 θ_1 和 θ_2 以及

$$\omega_{12}=d\tau-\pi^*\left(\frac{1}{q}\frac{\partial p}{\partial v}du+\frac{1}{p}\frac{\partial q}{\partial u}dv\right)$$

和

$$K=-\frac{1}{pq}\left[\frac{\partial}{\partial u}\left(\frac{1}{p}\frac{\partial q}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{1}{q}\frac{\partial p}{\partial v}\right)\right].$$

II.7 设 (u,v) 为习题 II.5 中那样的测地平行坐标系. 由 $C_u(v)=(u,v)$ 给出的曲线 C_u 是测地线. 因而 $\langle C''(v),\omega_{12}\rangle=0$. 但在我们的局部坐标下, 由于 $C'(v)$ 总是平行于基向量 f_i 中的一个, 并且因为沿 $C,u=$ 常数, 故

$$\langle C''(v),\widetilde{\pi}^*du\rangle=\langle C'(v),du\rangle=0,$$

从而 $\langle C''(v),d\tau\rangle=0$. 于是我们最后得到 $\frac{\partial q}{\partial u}=0$ 或者说 $q=q(v)$. 让我们把参数 v 换为 $w=\int_0^v q(t)dt$. 于是 (u,v) 是个测地平行坐标系, 对此坐标我们有

$$ds^2=pdu^2+qdw^2,$$

并且这时沿任意 $u=$ 常数的曲线的弧长为 $\int dw$.

II.8 证明当 $|w|$ 充分小时, 任意联结 (0,0) 和 $(0,w)$ 的曲线其弧长必定至少为 $|w|$.(由于我们原来的曲线 $x=0$ 的选取是随意的,) 我们因而有结论说, 测地线局部地极小化了长度.

II.9 设 $\prec w,z\succ$ 为一个黎曼流形中一个开集 U 上的局部坐标, 具有的性质是, 由 $C_z(w)=(w,z)$ 给出的曲线 C_z 是依弧长作参数的测地线. 因此 $z=$ 常数是条测地线且 $\left\|\frac{\partial}{\partial w}\right\|=1$. 设

$$\left(\frac{\partial}{\partial w},\frac{\partial}{\partial z}\right)=a.$$

证明 $\frac{\partial a}{\partial w}=0$.[提示: 由对 $\left(\frac{\partial}{\partial w},\frac{\partial}{\partial z}\right)$ 法正交化得到一个映射 ψ, 它的相关联的形式 θ_1 和 θ_2 弧由 $\theta_1=dw+adz,\theta_2=bdz$ 给出, 其中 $b=\left\|\frac{\partial}{\partial z}-a\left(\frac{\partial}{\partial w}\right)\right\|$. 然后计算 l_1 和 l_2, 并运用 $z=0$ 为测地线的事实.]

II.10 构造测地极坐标. 就是说, 对固定的 $p\in M$, 设 θ 为 $T_p(M)$ 中单位圆上的角坐标. 对每个 θ, 设 $C_\theta(\cdot)$ 为惟一的测地线, 它以弧长为参数, 并使得 $C(0)=p,C'(0)$ 对应于 $T_p(M)$ 中的角 θ. 证明映射 $(r,\theta)\longmapsto C_\theta(r)$ 给出了在 $U-\{p\}$ 上的"坐标", 其中 U 为 p 的某个邻域. 取 $\prec r,\theta\succ$ 为习题 II.9 中的 $\prec z,w\succ$ 并取 $r=0$ 的极限, 最后得到

$$\left(\frac{\partial}{\partial r},\frac{\partial}{\partial\theta}\right)\equiv 0.$$

因此利用 $U-\{p\}$ 上的"坐标" $\prec r,\theta\succ$, 其黎曼度量便有了形式

$$ds^2=dr^2+Ad\theta^2.$$

这些对于一般二维黎曼流形的坐标类比于第九章末尾中在平面和球面上引进的极坐标. 在那里所给出的论证一般地可用来给出测地线局部极小化了长度这个论断的另一个证明.

现在我们继续研讨等式

$$d\omega_{12}=-\pi^*(K\Omega)$$

的推论.

设 D 为 M 上具正则边界的区域, ψ 为 D 的某个邻域到 $\mathfrak{F}$ 的映射, 它满足 $\pi\circ\psi=$ 恒同映射. 于是由斯托克斯定理,

$$\int_{\partial D}\psi^*(\omega_{12})=-\int_D\psi^*\pi^*K\Omega=-\int_D K\Omega,\tag{II.12}$$

其中 Ω 为 M 上的面积形式.

让我们将此运用到下面构造的 ψ 上: 设 Y 为紧流形 M 上的一个向量场, 它只在有限个点 $p_1, \cdots, p_r$ 上为零. 在每个使 $Y_x \neq 0$ 的点 $x \in M$ 上, 设 $f_1(x) = \frac{Y_x}{\|Y_x\|}$, 令 $\psi(x) = \prec x, f_1(x), f_2(x) \succ$. 在每个点 p_i, 选取分图表 (U_i, h_i) 使得 U_i 不包含其他的点 p_j, 并在平面上有 $h_i(p_i) = (0,0)$. 设 $\gamma_{i,r}$ 为 $h_i^{-1}(C_r)$, 其中 C_r 是平面中半径为 r 的圆.[这里的 r 很小使得 C_r 在 $h_i(U_i)$ 内.] 设 $D = M - U_i(\gamma_{i,r}$ 的内部). 我们必须计算 $\int_{\gamma_{i,r}} \psi^*(\omega_{12})$, 其中的 $\gamma_{i,r}$ 是顺时针定向而不是逆时针的. 为此, 我们在 p_i 附近引进由坐标 x^1, x^2 来的法正交标架. 因此我们有了坐标 x^1, x^2, τ, 其中 $\tau(x, f_1, f_2)$ 度量了 $f_1(x)$ 与 $\frac{\partial}{\partial x^1}|_x$ 形成的角. 因此 (取 $\gamma_{i,r}, \tau$ 顺时针)

$$-\int_{\gamma_{i,r}} d\tau(f_1(s)) = 2\pi \cdot (Y\text{在 } p_i \text{ 的指数}).$$

但 $\psi^*\omega_{12} = \psi^* d\tau + \psi^*(k_1\alpha^1 + k_2\alpha^2)$, 故

$$-\int_{\gamma_{i,r}} \psi^*\omega_{12} = 2\pi \cdot (Y\text{在 } p_i \text{ 的指数}) + \int_{\gamma_{i,r}} k_1\theta^1 + k_2\theta^2.$$

现在让 $r \to 0$, 右端第二项为零, 而 (II.12) 右端有 $\int_D \to \int_M$. 因此我们便证明了下面的断言:

设 Y 为在有限个点 $p_1, \cdots, p_n$ 为零的向量场. 于是有

$$\sum \text{index}_{p_i}(Y) = \frac{1}{2\pi}\int_M K dA. \tag{II.13}$$

特别地, 一个向量场的指数和与此向量场无关, 同时曲率的全积分也与黎曼度量无关.

举例来说, 如果 M 为 S^2, 则切于经圆的向量场只在南北极为零, 而在这些点的指数为 $+1$. 因此 (II.13) 说 S^2 上任意向量场的指数和必为 2(特别地, 不可能在 S^2 上构造一个处处不为零的向量场). 另外, (II.13) 还说, 不管 S^2 上被赋予了什么黎曼度量, 我们总有

$$\int_{S^2} K dA = 4\pi.$$

相似地, 如果 M 为环面, 我们可在上面构造一个处处不为零的向量场, 因而 (II.13) 的那个整数为 0.

习题

设 M 表示环面, 角坐标为 φ_1, φ_2(准确到 2π). 映射 F 代表 M 到三维欧氏空间的一个浸入, 其定义为:

$$x_1 = (2+\cos\varphi_1)\cos\varphi_2,$$
$$x_2 = (2+\cos\varphi_1)\sin\varphi_2,$$
$$x_3 = \sin\varphi_1.$$

II.11 由 F 诱导的 M 上的黎曼度量是什么? 就是说, 对每个 $p \in M$ 计算 $\left(\frac{\partial}{\partial\varphi_1}, \frac{\partial}{\partial\varphi_1}\right)_p$, $\left(\frac{\partial}{\partial\varphi_1}, \frac{\partial}{\partial\varphi_2}\right)_p$ 和 $\left(\frac{\partial}{\partial\varphi_2}, \frac{\partial}{\partial\varphi_2}\right)_p$.(它们可以用 φ_1, φ_2 的函数给出.) 设 f_1, f_2 代表由对 $\frac{\partial}{\partial\varphi_1}, \frac{\partial}{\partial\varphi_2}$ 进行法正交化得到的向量场. f_1, f_2 的显式表达式是什么?

II.12 对于这个黎曼度量此环面的总面积为多少?

II.13 利用习题 II.11 给出的向量场 f_1, f_2 我们可以引进 $\mathfrak{F}$ 上的 (角) 坐标 $\varphi_1, \varphi_2, \tau$. 在这些坐标下, $\mathfrak{F}$ 上的向量值函数 X, f_1, f_2, f_3 的显式表达式是什么? (这些向量值函数应该用三维欧氏空间中固定的标准基给出, 即为 $\varphi_1, \varphi_2, \tau$ 的函数的三元组.) 因此

$$X(\varphi_1, \varphi_2, \tau) = \prec (2+\cos\varphi_1)\cos\varphi_2, (2+\cos\varphi_1)\sin\varphi_2, \sin\varphi_1 \succ.$$

对于 f_1, f_2, f_3 的相应表达式是什么?

II.14 由 $d\varphi_1, d\varphi_2, d\tau$ 给出的 $\mathfrak{F}$ 上的形式 $\omega_1, \omega_2, \omega_{12}$ 的显式表达式是什么?

II.15 M 的高斯曲率 K 是什么? (K 还是要作为 φ_1 和 φ_2 的函数来表示.) 在 M 上哪些点高斯曲率为正, 哪些点为负?

II.16 在任意的黎曼流形上有一种显然的方法把一个线性微分形式与向量场等同起来. (通过以 $T_x(M)$ 上的内积给出 $T_x^x(M)$ 与 $T_x(M)$ 之间的等同.) 如果 g 是个函数, 与形式 dg 相关联的向量场被记为 $\text{grad}g$. 设 M 为具上面所给的黎曼度量的环面, 并设 x_1 为由 $x_1(\varphi_1, \varphi_2) = (2+\cos\varphi_1)\cos\varphi_2$ 给出的函数. 证明向量场 $X = \text{grad}x_1$ 正好在四个点上为零, 计算这些点中每一个的指数.

设 x^1, x^2 为 $U \subset M$ 上的局部坐标系, $\psi: U \to \mathfrak{F}$ 由法正交化 $\prec \partial/\partial x^1, \partial/\partial x^2 \succ$ 给出的映射. 于是在 $\pi^{-1}(U)$ 上的坐标为 $x^1 \circ \pi, x^2 \circ \pi, \tau$, 其中 $\tau(z)$ 是 f_1 与 $\partial/\partial x^1$ 的夹角, 而 $z = \prec f_1, f_2 \succ$. 于是在 $\pi^{-1}(U)$ 上有

$$\omega_{12} = d\tau + k_1\alpha_1 + k_2\alpha_2 = d\tau + \pi^*\psi^*\omega_{12}.$$

设 C 为 U 中曲线, 满足 $C'(t)>0$, 而 $\widetilde{C}$ 为 $\widetilde{\pi}(U)$ 中曲线, 它对每个 t 派定了单位切向量 $C'(t)/\|C'(t)\|$. 于是

$$\int_{\widetilde{C}}\omega_{12}=\int\widetilde{C}^*\omega_{12}=\int_C\psi^*\omega_{12}+\int_C d\tau,$$

其中 τ 为 $C'(t)$ 与 x - 轴间的角.

从这个公式我们可以推导出许多有趣的结果, 我们将其以习题形式叙述于后.

II.17 证明一个测地三角形的内角和由 $2\pi-\int_D K\Omega$ 给出.(一个测地三角形是个由三条测地线为边界的区域.)

II.18 设 D 是个区域, 它在局部坐标下由一个简单多边形给出. 因此 $\partial D=C_1\cup\cdots\cup C_k$, 其中每条曲线 C_i 是在局部坐标系下的直线段. 设 $\alpha_1,\cdots,\alpha_k$ 为内角 (见图 11.18). 证明

$$k\pi+\int_D k\Omega=2\pi-\sum\int_{\widetilde{C}}\omega_{12}.$$

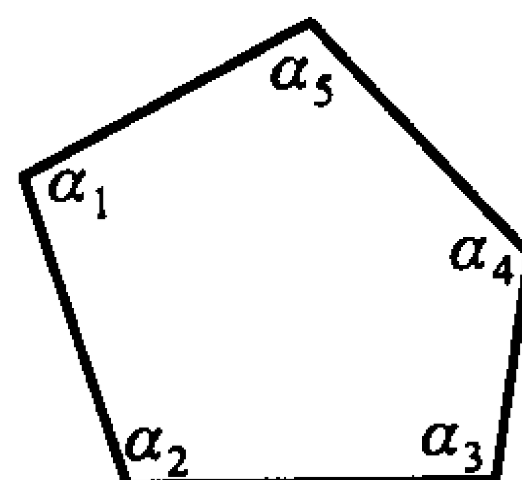


图11.18

II.19 如果 M 为紧曲面, 则此给了我们另一种计算 (II.13) 中整数的办法.

定义 II.1 一个紧二维微分流形 M 上的胞腔剖分是一个闭子集 $F_1,\cdots,F_m$ (称为胞腔) 的有限组, 使得 $M=\bigcup_{i=1}^m F_i$, 它们满足

(1) 对每个 F_i 存在 F_i 的一个邻域到一个具 n_i 条边的 $(n_i\geqslant 3)$ 的多边形上的一对一的双可微映射.

(2) 对 $i\neq j$, 或者 $F_i\cap F_j$ 为空或者 $f_i(F_i\cap F_j)$ 是对应多边形的一条固定边或一个顶点.

设 f 为此二维黎曼流形的胞腔剖分的面数, e 为边数, 而 v 为顶点数. 证明

$$f-e+v=\int_M K\Omega.$$

因此 $f-e+v$ 不依赖于此胞腔剖分. 因此我们给出了三种方法来计算与此流形相关的一个整数. 称这个整数为欧拉特征, 以 $\chi(M)$ 表示. 因此

$$\chi(M)=\int_M K\Omega=f-e+v=\sum\mathrm{index}_{p_i}Y$$

II.20 称一个胞腔剖分为一个正则的胞腔剖分是说每个面具有相同数目的边, 而每个顶点与相同数目的边相连. 证明在球面的正则胞腔剖分中最多只有五种可能的面数. 由此得出结论说, 最多只有五种“正立体”.

第十二章 $\mathbb{E}^n$ 中的位势理论

12.1 立体角

在下面 $x^1, \cdots, x^n$ 表示 $\mathbb{E}^n$ 中的欧几里得坐标. 设 $r^2 = (x^1)^2 + \cdots + (x^n)^2$; 于是 $dr^2 = 2rdr$, 故而我们有

$$\begin{aligned} rdr &= \sum x^i dx^i, \\ *rdr &= \sum (-1)^{i-1} x^i dx^1 \wedge \cdots \widehat{dx^i} \cdots \wedge dx^n, \\ d*rdr &= ndx^1 \wedge \cdots \wedge dx^n. \end{aligned}$$

设 i 表示单位球面的内射 $S^{n-1} \to \mathbb{E}^n$. 设 V_n 表示单位球的体积, A_{n-1} 表示 $(n-1)$ 维单位球面的体积. 半径为 r 的球面的体积便是 $r^{n-1}A_{n-1}$, 故

$$V_n = \int_0^1 r^{n-1} A_{n-1} dr = \frac{1}{n} A_{n-1}.$$

由于 $i^*(*rdr)$ 是一个 $(n-1)$ 形式, 它在旋转下不变, 所以它是 S^{n-1} 的体积形式的某个倍数. 由斯托克斯定理,

$$\int_{S^{n-1}} i^*(*rdr) = \int_{B^n} d*rdr = n \int_{B^n} dx^1 \wedge \cdots \wedge dx^n.$$

将此与上面的比较, 我们得到 $i^*(*rdr)$ 是单位球面上体积形式的结论.

设 ρ 代表 $\mathbb{E}^n - \{0\}$ 到单位球面上的投影. 令

$$\tau = \rho^* i^* (*rdr).$$

我们称 τ 为立体角元素. 它在任何一个 $(n-1)$ 维曲面上的积分给出了此曲面在 $(n-1)$ 维单位球面上投影的体积 (算上定向的符号和重数).

因为 $i^*(*rdr)$ 是在 $(n-1)$ 维流形 S^{n-1} 上的一个 $(n-1)$ - 形式, 而在 $(n-1)$ 维流形上的任意 $(n-1)$ - 形式的 d 必定为 0, 那么

$$d\tau = d\rho^* i^* (*rdr) = \rho^* (di^* (*rdr)) = 0,$$

所以我们有

$$d\tau = 0.$$

设 i_R 代表 S^{n-1} 作为半径为 R 的球面到 $\mathbb{E}^n$ 的内射 (故 $i_1 = i$). 于是直接从定义得到

$$i_R^* \tau = \frac{dS_R}{R^{n-1}},$$

其中 dS_R 为所诱导的体积形式 (见图 12.1).

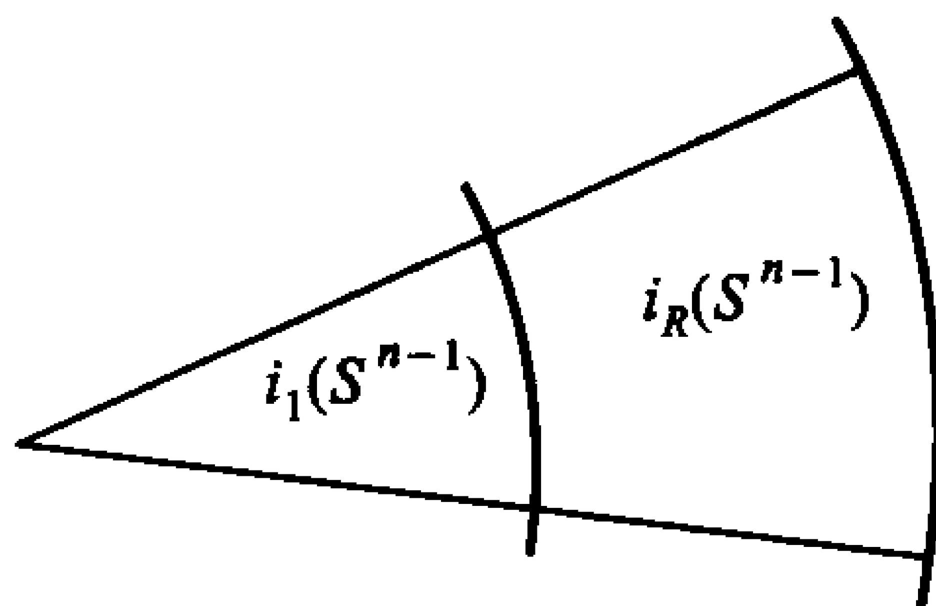


图 12.1

现在半径为 R 的球的体积为 $R^n V_n$, 半径为 R 的球面的体积 (面积) 为 $R^{n-1} A_{n-1} = R^{-1}(nR^n V_n)$. 这时 $i_R^*(*dr)$ 是个 $(n-1)$ - 形式, 因此有某个函数 f 使 $i_R^*(*dr) = f dS_R$. 由于 $*dr$ 和 dS_R 在旋转下都不变, 我们由此知道 f 是个常数. 但是

$$\int_{S^{n-1}} i_R^*(*rdr) = n \int_{B^n(R)} dx^1 \wedge \cdots \wedge dx^n = nR^n V_n,$$

故由此得到

$$dS_R = i_R^* \left(\frac{*rdr}{r} \right) = i_R^*(*dr).$$

因此

$$i_R^*\left(\tau - \frac{*dr}{r^{n-1}}\right) = 0.$$

现在设 x 为 $\mathbb{E}^n - \{0\}$ 中任意一个点, $\xi_1, \cdots, \xi_{n-1}$ 为 x 的切向量. 于是由上面的等式知道,

$$\left(\tau - \frac{*dr}{r^{n-1}}\right)(\xi_1, \cdots, \xi_{n-1})$$

在所有这些 ξ 切于中心在原点的球面, 并经过 x 时为零. 如果其中一个 ξ_i, 譬如 ξ_1 是 $\left(\frac{\partial}{\partial r}\right)_x$ 的倍数时, 这个表达式也为零, 其理由是 $\tau(\xi_1, \cdots, \xi_{n-1}) = i^*(*rdr)(\rho_*\xi_1, \cdots, \rho_*\xi_{n-1})$ 并且 $\rho_*\xi_1 = 0$, 以及 $*dr(\xi_1, \cdots, \cdot) \equiv 0$. 由多重线性性, 我们的结论是此表达式恒为零, 因此

$$\tau = \frac{*dr}{r^{n-1}} = \frac{*rdr}{r^n} = \frac{\sum(-1)^{i-1}x^i dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n}{r^n},$$

其中 $\wedge$ 表示对应的项已被略去.

12.2 格林公式

设 u 和 v 为 $\mathbb{E}^n$ 上的光滑函数. 于是

$$du = \sum \frac{\partial u}{\partial x^i} dx^i,$$

故

$$*du = \sum(-1)^{i-1}\frac{\partial u}{\partial x^i} dx^1 \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^n,$$

因此

$$d * du = \left[\sum \frac{\partial^2 u}{(\partial x^i)^2}\right] dx^1 \wedge \cdots \wedge dx^n = \Delta u dx^1 \wedge \cdots \wedge dx^n,$$

其中算子

$$\Delta = \frac{\partial^2}{(\partial x^1)^2} + \cdots + \frac{\partial^2}{(\partial x^n)^2}$$

被称作拉普拉斯算子.

设 U 是具紧闭包的开子集, 并有几乎正则的边界 (见图 12.2), 令

$$D_U[u, v] = \int_U du \wedge *dv = \int_U dv \wedge *du = \int \left(\sum_i \frac{\partial u}{\partial x^i}\frac{\partial v}{\partial x^j}\right) dx^1 \wedge \cdots \wedge dx^n.$$

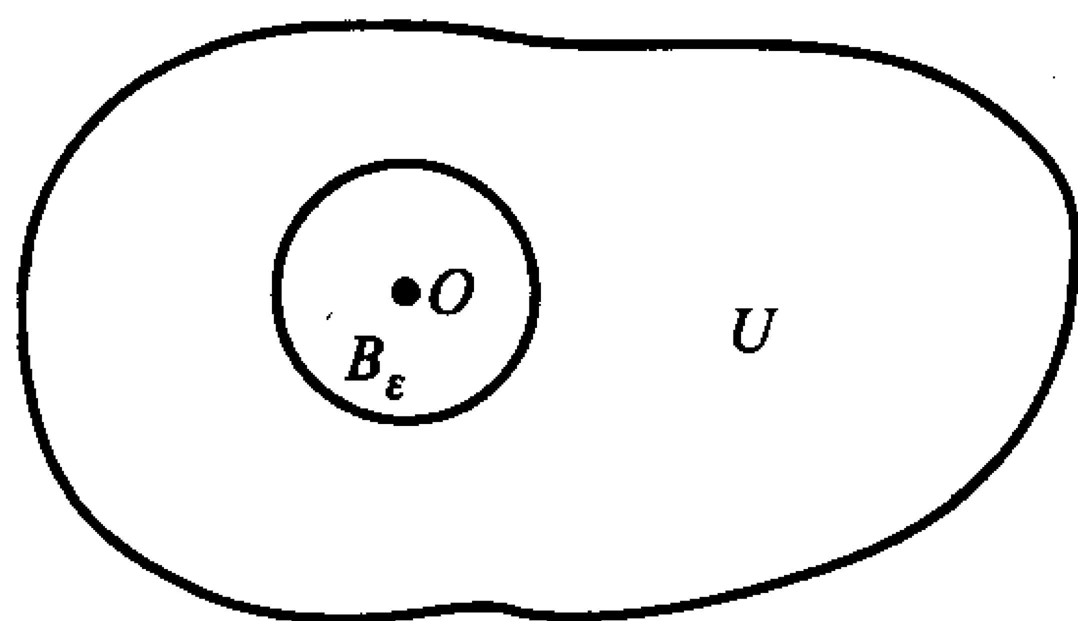


图 12.2

现在有 $\int_{\partial U} u * dv = \int_U d(u * dv)$. 但是 $d(u * dv) = du \wedge *dv + u \wedge d * dv$, 故

$$\int_{\partial U} u * dv = D_U[u, v] + \int_U (u\Delta v) dx^1 \wedge \cdots \wedge dx^n. \tag{2.1}$$

由于 D_U 对于 u 和 v 是对称的, 我们有

$$\int_{\partial U} u * dv - v * du = \int_U (u\Delta v - v\Delta u) dx^1 \wedge \cdots \wedge dx^n. \tag{2.2}$$

假设 U 包含了原点, 设 $U_\varepsilon = U - \overline{B_\varepsilon}$, 其中 B_ε 为中心在原点半径为 ε 的球. 设 $v = r^{2-n}$. [1] 则 $dv = (2-n)r^{1-n}dr$ 并且 $*dv = -(n-2)\tau$, 故 $d*dv$=0. 代入 (2.2) 式并取 ε 足够小使得 $\partial U_\varepsilon = \partial U - \partial B_\varepsilon$, 我们得到

$$-(n-2)\int_{\partial U}\left(u\tau + \frac{r^{2-n} * du}{n-2}\right) + (n-2)\int_{\partial B_\varepsilon} u\tau + \varepsilon^{n-2}\int_{\partial B_\varepsilon} *du$$
$$= -\int_{U_\varepsilon} r^{2-n}\Delta u dx^1 \wedge \cdots \wedge dx^n.$$

我们来考查一下当 $\varepsilon \to 0$ 时会发生什么. 左边第一项不依赖于 ε. 对第二项, 我们记 $u = u(0) + \mathfrak{O}(\varepsilon)$. 则第二项成了 $(n-2)A_{n-1}u(0) + \mathfrak{O}(\varepsilon)$. 对第三项, 我们写成

$$\int_{\partial B_\varepsilon} *du = \int_{B_\varepsilon} d * du = \int_{B_\varepsilon} \Delta u dx^1 \wedge \cdots \wedge dx^n.$$

因此左边第三项以 ε^{n-2} 阶成为 0. 由于 r^{2-n} 是 $\mathbb{E}^n$ 上的可积函数, 故右端趋向于 $-\int_U r^{2-n}\Delta u dx^1 \wedge \cdots \wedge dx^n$. 因此, 如果让 $\varepsilon \to 0$, 我们得到

$$A_{n-1}u(0) = \int_{\partial U}\left(u\tau + \frac{r^{2-n} * du}{n-2}\right) - \frac{1}{n-2}\int_U r^{2-n}\Delta u dx^1 \wedge \cdots \wedge dx^n. \tag{2.3}$$

[1] 对 $n = 2$, 令 $v = \log r$.

特别, 如果 u 是个调和函数, 即 $\Delta u = 0$, 则

$$u(0) = \frac{1}{A_{n-1}} \int_{\partial U} u\tau + \frac{1}{A_{n-1}(n-2)} \int_{\partial U} \frac{*du}{r^{n-2}}. \tag{2.4}$$

在 $U = B_a$ 的特殊情形, (2.4) 式右边第二项变成了

$$\frac{1}{(n-2)A_{n-1}a^{n-2}} \int_{\partial B_a} *du = \frac{1}{(n-2)A_{n-1}a^{n-2}} \int_{B_\varepsilon} d * du = 0.$$

因此 (由 A_{n-1} 和 τ 的定义) 有

$$u(0) = \frac{\int_{S_a} udS}{\int_{S_a} dS}, \tag{2.5}$$

其中 S_a 为半径为 a 的球面, dS 为其体积元. 换句话说, 如果 u 是个在某区域中的调和函数, 则 u 在任一点的值等于以此点为中心的任意球面上的平均值, 而此球面的内部完全包含在此区域内.

12.3 极大值原理

等式 (2.5) 有许多令人吃惊的结果, 我们现在来阐述它们. 设 u 是个在区域 U 中的调和函数. 设 x_0 为 U 中一个点, 并设存在 x_0 的邻域 W 使得 $W \subset U$, 并且对所有 $x \in W$,

$$u(x) \leqslant u(x_0).$$

(见图 12.3), 设 S_a 为半径为 a 中心在 x_0 的球面, 其中 a 小得使 $\overline{B}_a \subset W$. 于是对所有 $x \in S_a$ 有 $u(x) \leqslant u(x_0)$. 因此,

$$\int_{S_a} udS \leqslant u(x_0) \int_{S_a} dS.$$

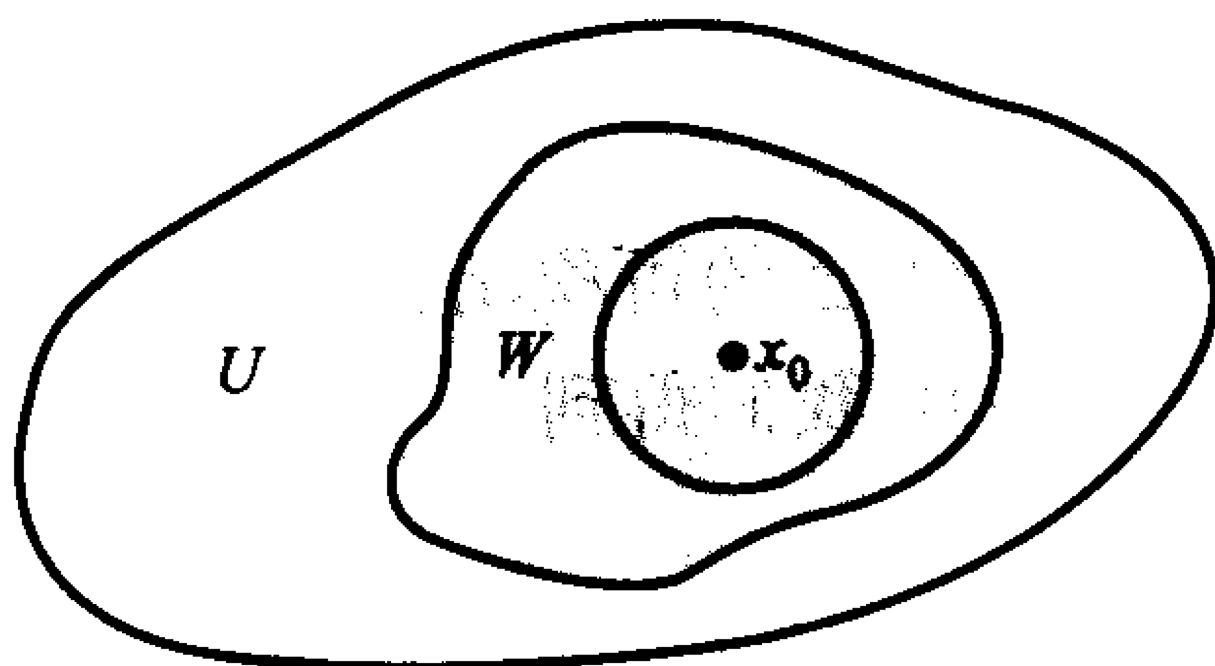


图 12.3

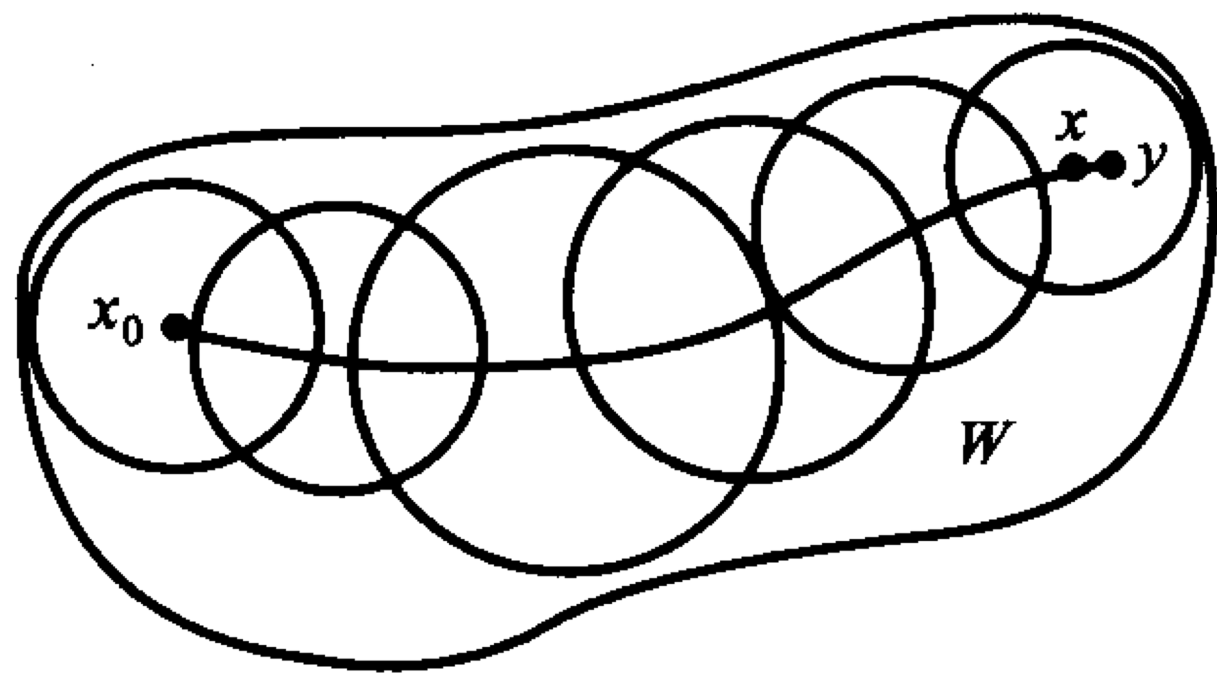


图 12.4

如果有某个点 $x \in S_a$ 使 $u(x) < u(x_0)$, 则由 u 的连续性知对所有充分靠近 x 的所有点 $y, u(y)$ 都小于 $u(x_0)$. 但是如此一来, 上面的不等式将会是严格地不相等, 即

$$\int_{S_a} udS < u(x_0)\int_{S_a} dS,$$

这与 (2.5) 式矛盾. 因此我们对所有 $x \in S_a$ 必有

$$u(x) = u(x_0).$$

现在假设 W 是个连通的开集, 即假定了 W 中任意两个点可以用全部在 W 中的一条连续曲线相联结. 设 y 为 W 中任一点, 而 C 为联结 x_0 和 y 的一条曲线. 在此曲线上每点 x 我们可以找到中心在 x 而整个在 W 内的一个充分小的球. 由 C 的紧性, 我们可以找到有限个这种球覆盖了 C. 因此可以正式叙述为: 存在有限个球面 $S_{a_1}, \cdots, S_{a_k}$, 使得每个球面及其内部都完全在 W 中, S_{a_1} 以 x_0 为中心, $S_{a_{i+1}}$ 的中心 x_i 在 S_{a_i} 上, 而 $y \in S_{a_k}$(见图 12.4). 但是这表明了 $u(x_0) = u(x_1) = \cdots = u(x_k) = u(y)$. 换句话说, 我们建立了下面的命题.

命题 3.1 设 u 在一个连通开集 W 内为调和函数, 并且假设 u 在某点 $x_0 \in W$ 取得了极大值. 于是, u 在 W 上为常数.

这个结果的一个直接推论是:

命题 3.2 设 U 为连通开集, 而 $\overline{U}$ 为紧. 则如果 u 是 $\overline{U}$ 上连续, U 中调和的函数, 则除了 u 是常数外有

$$u(x) < \max_{y \in \partial U} u(y). \tag{3.1}$$

证明 事实上, 因为 $\overline{U}$ 为紧, u 为连续, u 必定在 $\overline{U}$ 中某点 x_0 取得极大值. 如果我们可以实际上选取 $x_0 \in U$, 那么由命题 3.1, u 必定为常数. 如果 u 非常数, 则 $x_0 \in \partial U$, 从而有 (3.1) 式. □

由此我们推导出

命题 3.3 设 U 为连通的开集且 $\overline{U}$ 为紧. 设 u 和 v 为 U 中的调和函数并在 $\overline{U}$ 上连续. 假设对所有 $y \in \partial U$ 有

$$u(y) = v(y),$$

则对所有 $x \in U$ 成立

$$u(x) = v(x).$$

证明 事实上, $u-v$ 满足命题 3.2 的假设条件, 并在边界 ∂U 上为零. 因此对 $x \in U$ 有 $u(x)-v(x) \leqslant 0$. 相似地有 $v(x)-u(x) \leqslant 0$, 这蕴涵了此命题. □

另一种阐明命题 3.3 的方式是说, 在一个区域 U 上的调和函数完全被它的边界值决定. 这是一个惟一性定理: 最多只能有一个具已知的边界值的调和函数. 这提出了决定相对应的存在性定理是否成立的问题. 这就是有名的狄利克雷问题.

狄利克雷问题 已知定义于 ∂U 上的一个连续函数, 是否存在一个函数 u, 它在 $\overline{U}$ 上连续, 在 U 中调和, 并对所有 $y \in \partial U$ 有 $u(y) = f(y)$?

我们将在第 9 节中证明, 对于具有几乎正则边界的区域, 狄利克雷问题总是可解的.

12.4 格林函数

设 U 是个在上面可以解决狄利克雷问题的区域. 我们在这一节要证明这个解可以用在边界上的某个积分"显式地"给出这个解.

首先引进一些记号. 对每个 $x \in \mathbb{E}^n$ 令[2]

$$K_n(x,y) = \frac{1}{(n-2)A_{n-1}\|x-y\|^{n-2}}, \quad 其中 \quad y \in \mathbb{E}^n - \{x\}, (n \geqslant 3),$$

故 (对固定的 x) 在 $\mathbb{E}^n - \{x\}$ 上成立

$$*dK_n(x,\cdot) = \frac{1}{A_{n-1}}\tau_x,$$

其中的 τ_x 代表在点 x 的立体角. 于是对 $x \in U$, 等式 (2.3) 可以写为

$$u(x) = \int_{\partial U}[u * dK_n(x,\cdot) + K_n(x,\cdot) * du] - \int_U K_n(x,\cdot)\Delta U dV. \tag{4.1}$$

现在对固定的 $x \in U$, 函数 $K_n(x,y)$ 当 y 在 ∂U 上变化时是 y 的可微函数. 由假设条件, 我们因此可以找到一个函数 $h(x,\cdot)$, 它在 U 中调和, 在 $\overline{U}$ 上连续, 使得

$$h(x,y) = -K_n(x,y) \qquad 对所有 y \in \partial U 成立. \tag{4.2}$$

[2]这是对 $n \geqslant 3$ 的情形. 当 $n=2$ 时, 令 $K_2(x,y) = (1/2\pi)\ln|x-y|$.

另外, 对固定的 $y, h(x,y)$ 是 x 的连续函数. 事实上, 由极大值原理,

$$|h(x_1,y)-h(x_2,y)| \leqslant \max_{z\in\partial U}|K_n(x_2,z)-K_n(x_1,z)|.$$

而对所有 $z\in\partial U$, 只要 x 离 ∂U 维持一个固定的距离, $K_n(x,z)$ 便显然对 x 是一致连续的. 因此我们已经构造了一个函数 h_U 使得

(i) $h_U(x,y)$ 对 $x,y\in\overline{U}$ 是 x,y 的连续函数, 并对 $y\in U$ 对 y 可微;

(ii) 对每个固定的 $x, \Delta_y h_U=0$, 即 $\Delta h_U(x,\cdot)=0$;

(iii) $G_U(x,y)=K_n(x,y)+h_U(x,y)=0$ 对 $y\in\partial U$ 成立.

称函数 G_U 为区域 U 的**格林函数**. 我们暂且假定 G_U 存在, 并由此推导出它的一些性质. 记 $G=G_U$, 只要不会产生混淆我们就这样做. 首先要证明对 $x\in U, y\in U$, 我们有

$$G(x,y)=G(y,z). \tag{4.3}$$

设 $B_{x,\varepsilon}$ 和 $B_{y,\varepsilon}$ 为关于 x,y 的小球. 在 (2.2) 式中设 $u=G(x,\cdot), v=G(y,\cdot)$, 其中的 U 换做 $U-B_{x,\varepsilon}-B_{y,\varepsilon}$. 由于这两个函数都在此区域中为调和并在 ∂U 上为零, 我们得到

$$\begin{aligned}&\int_{\partial B_{x,\varepsilon}}G(x,\cdot)*dG(y,\cdot)+\int_{\partial B_{y,\varepsilon}}G(x,\cdot)*dG(y,\cdot)\\=&\int_{\partial B_{x,\varepsilon}}G(y,\cdot)*dG(x,\cdot)+\int_{\partial B_{y,\varepsilon}}G(y,\cdot)*dG(x,\cdot).\end{aligned}$$

我们要证明当 $\varepsilon\to0$ 时其左端趋向于 $G(x,y)$ 而右端趋向于 $G(y,x)$. 由对称性, 只看看左端就行了. 现在有

$$\begin{aligned}\int_{\partial B_{y,\varepsilon}}G(x,\cdot)*dG(y,\cdot)=&\int_{\partial B_{y,\varepsilon}}G(x,\cdot)*dK_n(y,\cdot)\\&+\int_{\partial B_{y,\varepsilon}}G(x,\cdot)*dh.\end{aligned}$$

由于 $A_{n-1}*dK$ 是关于 y 的立体角, 那么就像在 (2.3) 式的证明中那样, 右端第一项趋向于 $G(x,y)$. 又因为 $G(x,\cdot)$ 和 h 在 $B_{y,\varepsilon}$ 为光滑函数, 故第二项趋向于 0, 另一方面,

$$\int_{\partial B_{x,\varepsilon}}G(x,\cdot)*dG(y,\cdot)=\int_{\partial B_{x,\varepsilon}}K(x,\cdot)*dG(y,\cdot)+\int_{\partial B_{x,\varepsilon}}h(x,\cdot)*dG(y,\cdot).$$

如同上面那样, 第二项趋向于 0. 第一项 (对 $n \geqslant 3$) 可写为

$$\frac{1}{A_{n-1}(n-2)\varepsilon^{n-2}}\int_{\partial B_{x,\varepsilon}} *dG(y,\cdot) = \frac{1}{A_{n-1}(n-2)\varepsilon^{n-2}}\int_{B_{x,\varepsilon}} d*dG(y,\cdot) = 0,$$

这是因为 $G(y,\cdot)$ 在 $B_{x,\varepsilon}$ 为调和函数. 对于 $n=2$ 情形以 $\log\varepsilon$ 代替 ε^{n-2} 得到同样的论证.(4.3) 式得证.

设 u 为 U 上任意一个光滑函数. 将 (2.2) 式用于 $U - B_{x,\varepsilon}$ 上的 u 和 $v = G(x,\cdot)$. 我们得到 (由于在 ∂U 上 $G(x,\cdot)=0$)

$$\int_{\partial U} u*dG(x,\cdot) - \int_{\partial B_{x,\varepsilon}} u*dG(x,\cdot) + \int_{\partial B_{x,\varepsilon}} G(x,\cdot)*du$$
$$= \int_U G(x,\cdot)\Delta u dx^1 \wedge \cdots \wedge dx^n.$$

左端的第三个积分可以写为

$$\int_{\partial B_{x,\varepsilon}} K_n(x,\cdot)*du + \int h(x,\cdot)*du$$
$$= \frac{1}{A_n(n-2)\varepsilon^{n-2}}\int_{\partial B_{x,\varepsilon}} *du + \int_{\partial B_{x,\varepsilon}} h(x,\cdot)$$
$$= \frac{1}{A_n(n-2)\varepsilon^{n-2}}\mathfrak{O}(\varepsilon^{n-1}) + \mathfrak{O}(\varepsilon^{n-1})$$

从而趋向零.(对 $n-2$ 作通常的修改后同样可行.) 我们得到

$$u(x) = \int_U G(x,\cdot)\Delta u dx^1 \wedge \cdots \wedge dx^n + \int_{\partial U} u*dG(x,\cdot). \tag{4.4}$$

我们看到, (4.4) 式表明如果我们知道对 $\Delta F = f$ 存在一个满足边界条件 $F=0$ 的解, 则它由

$$F(x) = \int G(x,y)f(y)dy$$

给出. 相似地, 如果我们知道问题

$$\Delta u = 0, \qquad 且对 x \in \partial U 有 u(x) = f(x) \tag{4.5}$$

存在一个光滑解, 则它必由

$$u(x) = \int_{\partial U} u*dG(x,\cdot) \tag{4.6}$$

给出. 重要之处在于要看到这些公式是对 U 的格林函数的存在性的推论. 因此一旦我们能够找到函数 h 使其满足 (ii) 和 (iii), 则它们便都成立.

12.5 泊松积分公式

在这一节我们要对半径 R 的球显式地构造格林函数. 设 B_R 为半径为 R, 中心在原点的球.

对任意 $x \in \mathbb{E}^n - \{0\}$ 设 x' 为 x 关于半径为 R 的球面的反演映射的像:

$$x' = \frac{R^2}{\| x \|^2} x.$$

定义函数 G_R 为[3]

$$A_{n-1}(n-2)G_R(x,y) = \begin{cases} \dfrac{1}{\| y-x \|^{n-2}} - \dfrac{R^{n-2}}{\| x \|^{n-2} \| y-x' \|^{n-2}}, & \text{如果} x \neq 0, \\ \dfrac{1}{\| y \|^{n-2}} - \dfrac{1}{R^{n-2}}, & \text{如果} x = 0. \end{cases} \tag{5.1}$$

如果 $x \in B_R$ 到 $x' \notin B_R$, 故 (5.1) 式右端的第二项为 $\overline{B}_R$ 上的连续调和函数. 我们仅仅要验证性质 (iii) 成立. 现在对 $\| y \| = R$, 我们由相似三角形 (或直接计算) 有

$$\frac{R}{\| x \|} = \frac{\| y - x' \|}{\| y - x \|}, \tag{5.2}$$

使得对 $\| y \| = R$ 有

$$G_R(x,y) = 0$$

(见图 12.5). 这就是 (iii), 从而我们已经验证了 G_R 对于半径为 R 的球是其格林函数.

要应用 (4.4) 式我们必须在半径为 R 的球面上计算 $*dG_R$. 现在用 (5.1) 式我们 (对 $x \neq 0$) 有

$$\begin{aligned} A_{n-1} * dG_R(x, \cdot) &= \tau_x - \frac{R^{n-2}}{\| x \|^{n-2}} \tau_{x'} \\ &= \sum \left\{ \frac{y^i - x^i}{\| y - x \|^n} - \frac{R^{n-2}}{\| x \|^{n-2}} \frac{y^i - x^i}{\| y - x' \|^n} \right\} * dy^i. \end{aligned}$$

---

[3]如果 $n = 2$, 令 $G_R(x,y) = \begin{cases} \log \| y - x \| - \log \dfrac{\| x \|}{R} \| y - x' \|, & \text{如果} x \neq 0, \\ \log \| y \| - \log R, & \text{如果} x = 0. \end{cases}$

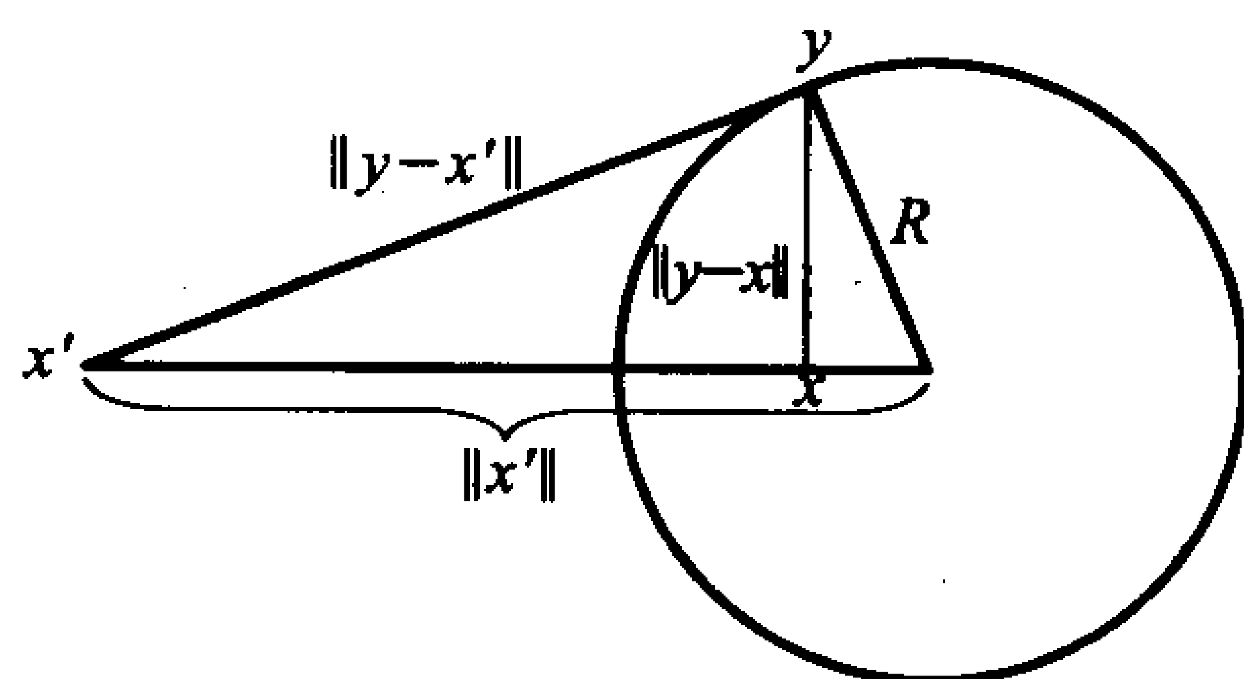


图 12.5

但由 (5.2) 式, 如果 $\| y \|= R$, 则

$$\frac{R^{n-2}}{\| x \|^{n-2} \| y - x' \|^n} = \frac{\| x \|^2}{\mathbb{R}^2 \| y - x \|^n}.$$

因此我们看到, 对 $\| y \|= R$ 有

$$\begin{aligned} &i_R^*(A_{n-1} * dG_R) \\ &= i_R^* \left\{ \frac{1}{\| y - x \|^n} \sum \left[y^i - x^i - \frac{\| x \|^2}{R^2} \left(y^i - \frac{R^2}{\| x \|^2} x^i \right) \right] * dy^i \right\} \\ &= i_R^* \left\{ \frac{R^2 - \| x \|^2}{\| y - x \|^n R^2} \sum y^i * dy^i \right\} \\ &= i_R^* \left\{ \frac{R^2 - \| x \|^2}{R^2 \| y - x \|^n} * rdr \right\}. \end{aligned}$$

但 $i_R^*(*rdr) = RdS_R$, 其中的 dS_R 是半径为 R 的球面的体积元. 如果将其代入 (4.6) 式, 我们得到了

$$u(x) = \frac{R^2 - \| x \|^2}{RA_{n-1}} \int_{S_R} \frac{u(y)}{\| y - x \|^n} dS_R \quad (\text{泊松积分公式}). \tag{5.3}$$

在 (5.3) 式的证明中我们用了一些假设条件, 即假定函数 u 在球 $\overline{B}_R$ 的某个邻域中可微, 并在 $\| x \|< R$ 时为调和. 实际上, 所有我们需要假定的是 u 对 $\| x \|< R$ 为可微且调和, 并在闭球 $\| x \|\leqslant R$ 上连续. 事实上, 对任意 $\| x \|< R$, 等式 (5.3) 在以 R 换作 R_a 时成立, 其中 $\| x \|< R_a < R$. 然后我们让 R_a 趋向于 R, 我们便利用所假定的 u 的连续性恢复了 (5.3) 式的有效性.

等式 (5.3) 给出了一个球的狄利克雷问题解是在假定我们知道了这个解存在的条件下才成立的. 那就是说, 如果 u 是任何一个在开球上调和在闭球上连续, 那么它才满足 (5.3) 式. 现在我们要证明 (5.3) 式实际上是对所指定的边值的

狄利克雷问题的一个解. 那么, 假设我们已知定义于球面 S_R 上一个连续函数 u. 我们对所有 $y \in S_R$ 有了 $u(y)$. 对 $\| x \| < R$ 用 (5.3) 式定义 $u(x)$. 我们必须证明

(a) u 对 $\| x \| < R$ 是调和的, 而且

(b) 如果 $x \longrightarrow y_0, \| y_0 \| = R$, 则 $u(x) \longrightarrow u(y_0)$.

要证明 (a), 我们注意到 $G_R(x,y)$ 是在 $\| x \| < R_1 < R, R_1 < \| y \| < \frac{R_2}{R_1}$ 中的 x 和 y 的可微函数, 并且由构造知道, 它还是 y 的调和函数. 对 $\| x \| < R, \| y \| < R$, 由 (4.4) 式我们知道有

$$G_R(x,y) = G_R(y,x).$$

因此对 $\| y \| < R$ 中固定的 $y, G_R(\cdot,y)$ 在 $B_R - \{y\}$ 上是个调和函数. 让 $\| y \| \longrightarrow R$, 我们看出 $G_R(x,y)$ 对于 $\| x \| < R_1 < R$ 和对于每个固定的 $y \in S_R$, 它是 x 的调和函数. 因而

$$\frac{\partial G_R}{\partial y^i}(x,y)$$

对每个 $y \in S_R$ 是 x 的调和函数. 换句话说, $*dG_R(\cdot,y)$ 的所有系数对每个 $y \in S_R$ 都是 x 的调和函数, 因此 $u(y) * dG_R(\cdot,y)$ 的每个系数也都如此. 由此得到函数 $\int_{S_R} u * dG_R(x,\cdot)$ 是 x 的调和函数, 这是因为这个积分 (相当于做关于 x 的各种导数的积分) 对 $\|x\| < R_1 < R$ 是一致收敛的.(a) 得证.

要证 (b), 我们首先注意常数 1 是个处处调和的函数, 故等式 (5.3) 可用于它. 因此我们有

$$\frac{R^2 - \|x\|^2}{A_{n-1}R} \int_{S_R} \frac{dS_R}{\|y-x\|^n} = 1 \qquad \text{对任意}\|x\| < R\text{成立}. \tag{5.4}$$

现设 y_0 为 S_R 中一个点, u 为 S_R 上一个连续函数. 对任意 $\varepsilon > 0$ 可以找到一个 $\delta > 0$ 使得当 $\|y - y_0\| < 2\delta (y \in S_R)$ 有

$$|u(y) - u(y_0)| < \varepsilon.$$

设 $Z_1 = \{y \in S_R : \|y - y_0\| > 2\delta\}, Z_2 = \{y \in S_R : \|y - y_0\| \leqslant 2\delta\}$. 于是由 (5.3) 和 (5.4) 式, 对 $\|x\| < R$ 我们有

$$u(x) - u(y_0) = \frac{R^2 - \|x\|^2}{(A_{n-1})R} \int_{S_R} \frac{u(y) - u(y_0)}{\|y-x\|^n} dS_R,$$

所以

$$|u(x) - u(y_0)| \leqslant I_1 + I_2,$$

其中

$$I_1 = \frac{R^2 - \|x\|^2}{A_{n-1}R}\int_{Z_1}\frac{|u(y)-u(y_0)|}{\|y-x\|^n}dS_R,$$
$$I_2 = \frac{R^2 - \|x\|^2}{A_{n-1}R}\int_{Z_2}\frac{|u(y)-u(y_0)|}{\|y-x\|^n}dS_R.$$

现在如果 $\|y_0 - x\| < \delta$, 则对所有 $y \in Z_1$, 我们有 $\|y-x\| > \|y-y_0\| - \|x-y_0\|$, 所以 $\|y-x\| > \delta$; 因此对所有使 $\|x-y_0\| < \delta$ 的 x, 在 I_1 中的积分为一致有界. 因为当 $x \to y_0$ 时 $\|x\| \to R$, 我们最后有当 $x \to y_0$ 时 $I_1 \to 0$.

关于 I_2, 我们知道对所有 $y \in Z_2$ 有 $|u(y)-u(y_0)| < \varepsilon$, 所以由 (5.4) 式有

$$I_2 < \frac{R^2 - \|x\|^2}{A_{n-1}R}\int_{Z_2}\frac{\varepsilon dS_R}{\|y-x\|^n}$$
$$< \varepsilon\left(\frac{R^2 - \|x\|^2}{A_{n-1}R}\int_{S_R}\frac{dS_R}{\|y-x\|}\right) = \varepsilon$$

证明了 (b).

因此我们便证明了:

定理 5.1 设 u 为定义在球面 S_R 上的连续函数. 存在定义于 $\|x\| < R$ 上的一个惟一的连续函数, 它在球面 S_R 上与已知函数 u 一致, 并且对 $\|x\| < R$ 为调和. 这个函数由 (5.3) 式对所有 $\|x\| < R$ 给出.

12.6 泊松积分公式的推论

在定理 5.1 的证明中我们假定了 u 在那个闭球中连续, 而在开球中有两个连续的导数, 并在开球中调和. 然而从 (5.3) 式可以清楚看到, u 对 $\|x\| < R$ 具有任意阶的导数. 事实上, 如果 $\|x\| < R_1 < R$, 我们可以微分 (5.3) 式的积分号内任意次, 这是因为我们所得到的所有积分对 $\|x\|$ 是一致可积的. 现在如果 u 在某个开集 U 内调和, $x \in U$, 我们则可选取 R 充分小, 使得半径为 R, 中心在 x 的球包含在 U 内 (见图 12.6). 于是我们可应用 (5.3) 式. 我们的结论是:

命题 6.1 设 u 为定义在开集 U 上的函数, 它具有两个连续导数并满足 $\Delta u \equiv 0$. 于是 u 具有所有阶的连续导数.

只要更仔细考查 (5.3) 式, 我们还可以改进命题 6.1. 设 $\|y\| = R, R_1 < R$. 因

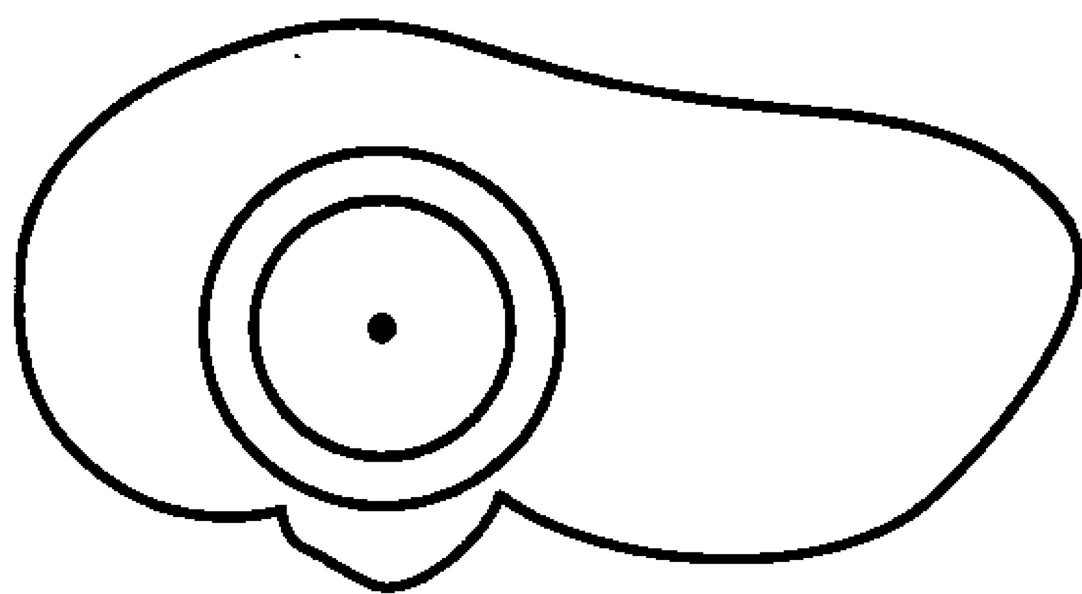


图 12.6

此所有 $\|x\| \leqslant R_1$. 多项式定理让我们展开

$$\frac{1}{\|y-x\|^n} = \sum A_{i_1,\cdots,i_n}(y)x_1^{i_1}\cdots x_n^{i_n}, \tag{6.1}$$

其中的系数依赖于 y 但此级数对所有 $\|y\| = R$, 关于 x 一致收敛. 另外, 如果 D 是关于 x 的任意一个偏微分算子, 我们则得到一个相类比的展开式

$$D(\|y-x\|^{-n}) = \sum B_{\alpha_1,\cdots,\alpha_n}(y)x_1^{\alpha_1}\cdots x_n^{\alpha_n},$$

其中的这个级数是由 (5.4) 式逐项微分并在此同一区域中一致收敛得到的. 如果我们把 (6.1) 式代入 (5.3) 式, 则因为对 y 的一致收敛性, 可以逐项积分此级数, 最后得到

$$u(x) = \sum C_{\alpha_1,\cdots,\alpha_n}x_1^{\alpha_1}\cdots x_n^{\alpha_n},$$

这里的这个级数对 $\|x\| < R_1$ 一致收敛. 另外, 我们可以逐项微分此级数. 在这样做时, 取在 $x=0$ 的值, 我们便看到

$$\alpha! C_\alpha = D^\alpha u(0),$$

其中 $\alpha = \prec \alpha_1, \cdots, \alpha_n \succ$.

因此我们证明了:

命题 6.2 *设 u 在球 $\|x\| < R$ 中为调和. 于是*

$$u(x) = \sum \frac{1}{\alpha!} D^\alpha u(0) x^\alpha, \tag{6.2}$$

其中的泰勒级数 (6.2) 对所有 $\|x\| < R$ 收敛, 且对所有 $\|x\| \leqslant R_1 < R$ 一致收敛. 特别, 在整个球上 u 由它的值和它在原点的所有它的导数的值所决定.

设 U 为 $\mathbb{E}^n$ 中某个连通的开子集, 并设 u 和 v 为 U 上的两个调和函数. 假设 u 和 v 在某点 $x \in U$ 有相同的值和相同的所有阶的导数值. 设 y 为 U 中某

个其他点. 我们于是可以用一系列 U 中的球联结 x 和 y, 其中每个球的中心在前面一个球的内部, 并使 x 为第一个球的中心, y 则在最后一个球内 (见图 12.7). 因此我们有结论

$$u(y) = v(y).$$

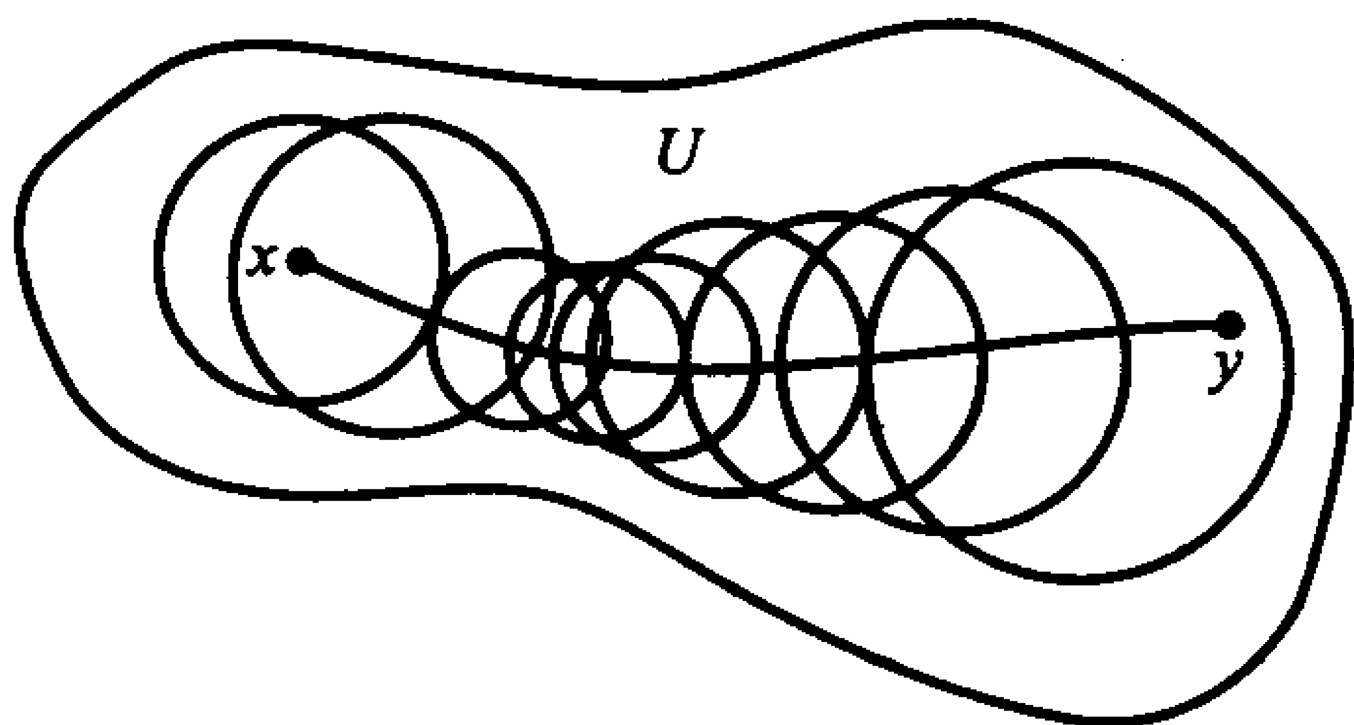


图 12.7

从而我们有:

命题 6.3 设 u 和 v 为定义在开连通集 U 上的调和函数. 假设 u 和 v 连同它们所有的导数在 U 中某个点相同. 则在 U 上 $u \equiv v$. 特别地, 如果 $u(x) = v(x)$ 对所有 $x \in W$ 成立, 其中 W 为 U 中某个开子集, 则对所有 $x \in U$ 有 $u(x) = v(x)$.

我们继续考查 (5.3) 式的推论. 设 u 由 (5.3) 式给出. 设 D 表示任何关于 x 的偏微分算子. 因而

$$D^\alpha u = \frac{\partial^{\alpha_1+\cdots+\alpha_n}}{(\partial x^1)^{\alpha_1}\cdots(\partial x^n)^{\alpha_n}}.$$

于是 $D^\alpha u(x)$ 由一个在 S_R 上的积分给出, 它的被积函数最坏也只包含了 $\|y-x\|$ 的逆, 其中 y 在 S_R 上, 和在 S_R 上的函数 u. 特别地, 如果 $\|x\| < R_1 < R$, 我们则可用 u 在 S_R 上的值和差值 $R - R_1$ 去估计 $D^\alpha u$ 的极大绝对值. 简单地说,

$$|D^\alpha u(x)| \leqslant c(D^\alpha, R, R_1) \max_{\|y\|=R} |u(y)|, \text{其中} \|x\| < R_1, \tag{6.3}$$

这里的 c 只依赖于 α, R 和 R_1. 现在假设 u 在某个开集 U 中为调和, K_1 和 K_2 是 U 中的紧子集, 满足

$$K_1 \subset \mathrm{int}K_2 \subset K_2 \subset U.$$

在每个 $x \in K_1$ 附近我们可以作一个开球 B_{R_x} 使得 $B_{R_x} \subset \mathrm{int}K_2$, 从而

$$S_{R_x} \subset K_2.$$

我们也可作一个 x 附近的一个具稍小半径的球 $B_{R_{1x}}$. 现在 $B_{R_{1x}}$ 覆盖了 K_1. 因为 K_1 为紧, 我们可选取有限个这样的球覆盖 K_1. 把 (6.3) 式用到每个这种球上, 并把 $|u(y)|$ 换成较大的 $\max_{z\in k_2}|u(z)|$, 再用有限多个常数 c 中最小的一个, 我们最终得到

$$\max_{z\in K_1}|D^\alpha u(x)| < c(\alpha, K_1, K_2)\max_{z\in K_2}|u(z)|. \tag{6.4}$$

特别, 我们有

命题 6.4 设 $\{u_k\}$ 为定义在一个开集 U 上的调和函数序列, 它在 U 的任一紧子集上一致收敛. 于是偏导数的序列 $\{D^\alpha u_k\}$ 也在任意紧子集上一致收敛. 特别, 这个序列的极限仍然是调和函数.

事实上, 对任意紧子集 K_1, 我们总能找到一个紧子集 K_2 使得 $K_1\subset \mathrm{int}K_2$. 把 (6.4) 式用于调和函数 u_i-u_j 便建立了在 K_1 上 $\{D^\alpha u_k\}$ 的一致收敛性. 因为偏导数的序列一致收敛, 所以 $\Delta u=\lim\Delta u_k=0$.

12.7 哈纳克定理

我们继续收获等式 (5.3) 的成果. 除了在上一节中我们所假定了的那些条件外, 我们还假设对所有 $y\in S_R$ 有 $u(y)\geqslant 0$.

现在对 $\|x\|=R_1<R$ 我们有 (见图 12.8)

$$R-R_1\leqslant\|y-x\|\leqslant R+R_1 \qquad \text{对所有}\|y\|=R\text{成立}.$$

于是由 (5.3) 式,

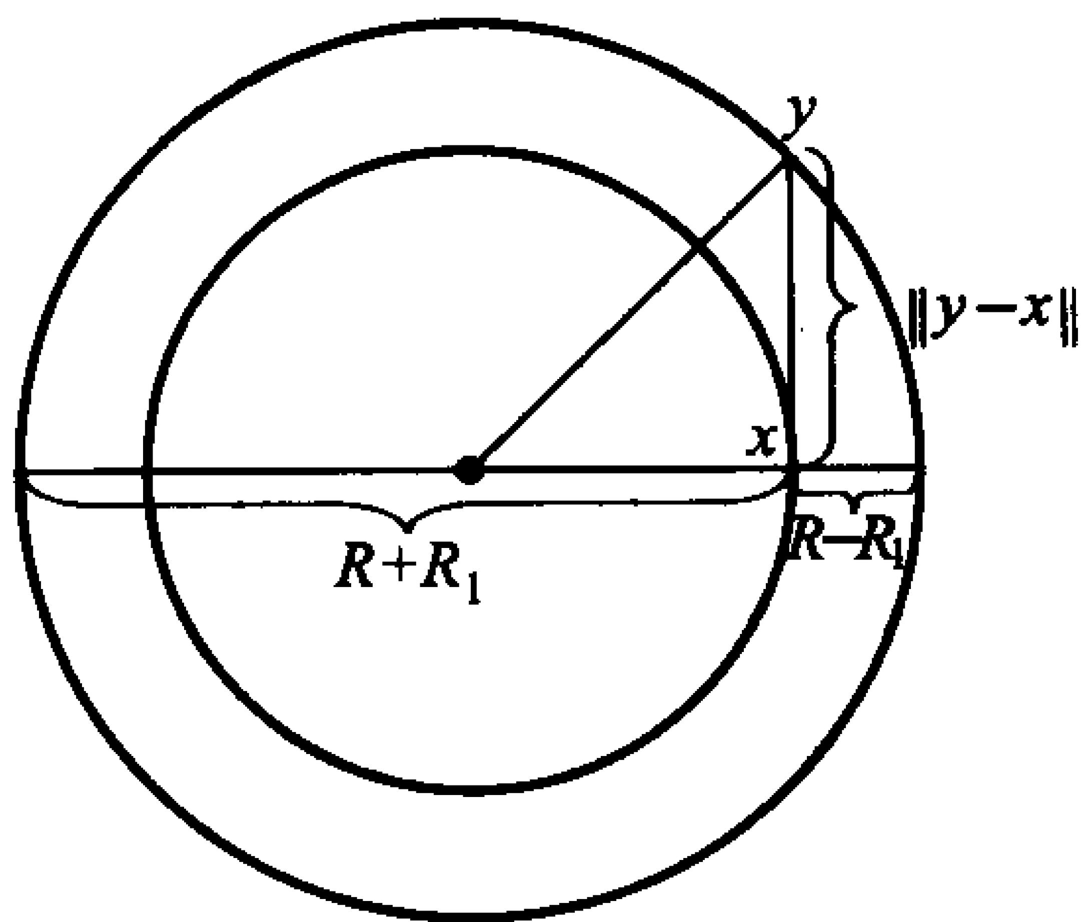


图 12.8

$$\frac{R^2-R_1^2}{A_{n-1}R(R+R_1)^n}\int_{S_R} udS_R \leqslant u(x) \leqslant \frac{R^2-R_1^2}{A_{n-1}R(R-R_1)^n}\int_{S_R} udS_R.$$

根据 (2.5) 式, 在此不等式右边和左边出现的积分等于 $A_{n-1}R^{n-1}u(0)$. 因此

$$\frac{(R^2-R_1^2)R^{n-2}}{(R+R_1)^n}u(0) \leqslant u(x) \leqslant \frac{(R^2-R_1^2)R^{n-2}}{(R-R_1)^n}u(0). \tag{7.1}$$

不等式 (7.1) 以哈纳克不等式而知名. 它有下面的推论. 假设 $\{u_k\}$ 是满足 (5.3) 式的函数序列并且使得

$$u_1(y) \leqslant u_2(y) \leqslant \cdots \leqslant u_k(y) \leqslant \cdots \qquad 对所有y \in S_R成立.$$

如果我们把 (7.1) 式用于函数 $u_j - u_i (j \geqslant i)$, 我们便得到

$$|u_j(x) - u_i(x)| \leqslant d(R, R_1)|u_j(0) - u_i(0)| \qquad 对所有\|x\| \leqslant R_1成立,$$

其中 $d(R, R_1)$ 只依赖于 R 和 R_1.

特别地, 如果此序列在原点收敛, 则它对 $\|x\| \leqslant R_1 < R$ 一致收敛. 运用我们通常使用的技巧, 用一系列的球联结两个点, 我们便推导出

命题 7.1 (哈纳克定理) *设 $\{u_k\}$ 为定义于一个连通开集 U 上的调和函数序列, 使得*

$$u_1(x) \leqslant u_2(x) \leqslant \cdots \qquad 对所有x \in U成立. \tag{7.2}$$

假设序列 $\{u_k(p)\}$ 对某个 $p \in U$ 收敛. 则函数序列 $\{u_k\}$ 在 U 的任何一个紧子集上一致收敛, 并且由命题 6.4 知, 它的极限函数仍是调和的.

哈纳克定理的一个有用的形式是:

命题 7.2 *设 $\{u_k\}$ 是定义在开集 U 上的调和函数序列并满足 (7.2) 式. 假设存在常数 M 使得 $u_k(x) < M$ 对所有 k 和所有 $x \in U$ 成立. 于是函数序列 $\{u_k\}$ 在 U 的任一个紧子集上一致收敛于一个调和函数.*

要证明命题 7.2, 我们注意到这一次在每个 $x \in U, \{u_k(x)\}$ 的收敛性是自动的, 因为它是有界实数的单调 (不降) 序列. 命题 7.1 保证了在紧子集上一致收敛到调和极限函数.

12.8 次调和函数

在这一节和下一节我们要证明对于边界满足某些正则性条件的有界开集 U, 狄利克雷问题能够有解. 我们所提出的证明 (有许多其他的证明) 属于佩隆 (Perron), 而且对于次调和函数的概念作了实质性的应用. 设 u 是定义在开集 U 上的函数. 我们称 u 是*次调和*的是说

(a) u 为连续; 以及

(b) 对任意 U 的连通开子集 W 和任何定义于 W 上的调和函数 v, 函数 $u-v$ 在 W 上满足极大值原理. 换句话说, 对任何这样的 W 和 v, 如果有某个 $x_0 \in W$, 对所有 $x \in W$ 满足

$$u(x) - v(x) \leqslant u(x_0) - v(x_0)$$

则在 W 中有 $u(x) - v(x) \equiv u(x_0) - v(x_0)$.

为了更好一点理解条件 (b), 我们来研究它的一些推论. 首先, 我们可以设 v 为恒等于 0 的调和函数. 于是 (b) 表明 u 在 U 的每个开子集上满足极大值原理.

其次, 设 B_R 为中心在 z, 闭包在 U 中的某个球. 特别地, 它的边界 S_R 包含在 U 中, 而函数 u 在 S_R 上连续. 我们因此能够找到一个定义于此开球上的调和函数 v, 它对 $y \in S_R$ 取值 $u(y)$. 取 W 为此开球, v 为刚刚构造的这个函数. 于是 $u - v$ 在 S_R 上为零, 故 (b) 意味着对所有 $x \in W$ 有 $u(x) \leqslant v(x)$. 特别, $u(z) \leqslant v(z)$. 但

$$v(z) = \frac{1}{A_{n-1}R^{n-1}} \int_{S_R} u dS_R.$$

我们从而证明了如果 u 是定义在 U 上的次调和函数, S_R 为中心在 z 的 U 中的球面, 则

$$u(z) \leqslant \frac{1}{A_{n-1}R^{n-1}} \int_{S_R} u dS_R.$$

换句话说, 一个次调和函数总是小于或等于它在以此点为中心的球面上的平均值. 这个性质常常被当作一个函数为次调和的定义, 而连续性的假定则多少放宽了. 然而我们在上面所给的定义对我们当前的目的而言更为适合.

设 w_1 和 w_2 为两个次调和函数, 定义于开集 U. 定义函数 $w_1 \vee w_2$ 为

$$(w_1 \vee w_2)(x) = \max[w_1(x), w_2(x)].$$

函数 $w_1 \vee w_2$ 仍是次调和的. $w_1 \vee w_2$ 为连续可由下面得到: 设 x 为 U 中任意点, 则 $w_1 \vee w_2(x)$ 或是 $w_1(x)$ 或是 $w_2(x)$, 譬如是 $w_1(x)$. 因为 w_1 为连续, 故对

任意 $\varepsilon > 0$ 我们可以找到 $\delta > 0$ 使得当 $|y-x| < \delta$ 时 $|w_1(y) - w_1(x)| < \varepsilon$. 相似地, 我们可对 y 在某个范围内使 $w_2(y) < w_2(x) + \varepsilon$, 故 $w_2(y) < w_1(x) + \varepsilon$(因为 $w_1(x) = \max[w_1(x), w_2(x)]$). 对这些 y 的值我们便有

$$w_1(x) - \varepsilon < w_1(y) \leqslant w_1 \vee w_2(y) < w_1(x) + \varepsilon,$$

故 $w_1 \vee w_2$ 连续. 现在设 W 是 U 的一个连通开子集, v 是 W 上的一个调和函数. 假设 $w_1 \vee w_2 - v$ 在 W 中某点 x_0 取得极大值. 假设 $w_1 \vee w_2(x_0) = w_1(x_0)$. 对于所有 $x \in W$ 则有

$$w_1(x) - v(x) \leqslant w_1 \vee w_2(x) - v(x) \leqslant w_1(x_0) - v(x_0).$$

由于 w_1 是次调和的, 这个不等式的左右两端必相等. 因此 $w_1 \vee w_2 - v$ 在 W 上满足极大值原理, 从而证明了 $w_1 \vee w_2$ 是次调和的.

设 w 为定义在开集 U 上的次调和函数, B 为闭包在 U 中的一个球. 设 u 为对球 B 的边界值为 w 的狄利克雷问题的解. 因此 u 是惟一在闭球中连续, 在开球中调和而在 B 的边界 S 上等于 w 的函数. 如同我们已经看到的, 对 $x \in B$ 有 $w(x) \leqslant u(x)$. 现在定义函数 w_B 为

$$w_B(x) = \begin{cases} w(x), & \text{当} x \in U - B, \\ u(x), & \text{当} x \in B. \end{cases}$$

我们断言函数 w_B 又是一个次调和函数. 证明同样如前. 连续性像上面一样是显然的. 如果 W 是 U 的子集而 v 在 W 上调和函数, 则 $w_B - v$ 在任何内点都不能取得严格的极大: 假设对所有 $x \in W$ 我们有 $w_B(x) - v(x) \leqslant w_B(x_0) - v(x_0)$(见图 12.9). 如果 $x_0 \in U - B$, 我们则有 $w(x) - v(x) \leqslant w_B(x) - v(x) \leqslant w_B(x_0) - v(x_0) = w(x_0) - v(x_0)$. 因为 w 是次调和的, 所有这些不等号都应相等. 另一方面, 如果 $x_0 \in B$, 则由于 w_B 在开球内是调和的, 我们必有 $w_B(x) - v(x) = w_B(x_0) - v(x_0)$

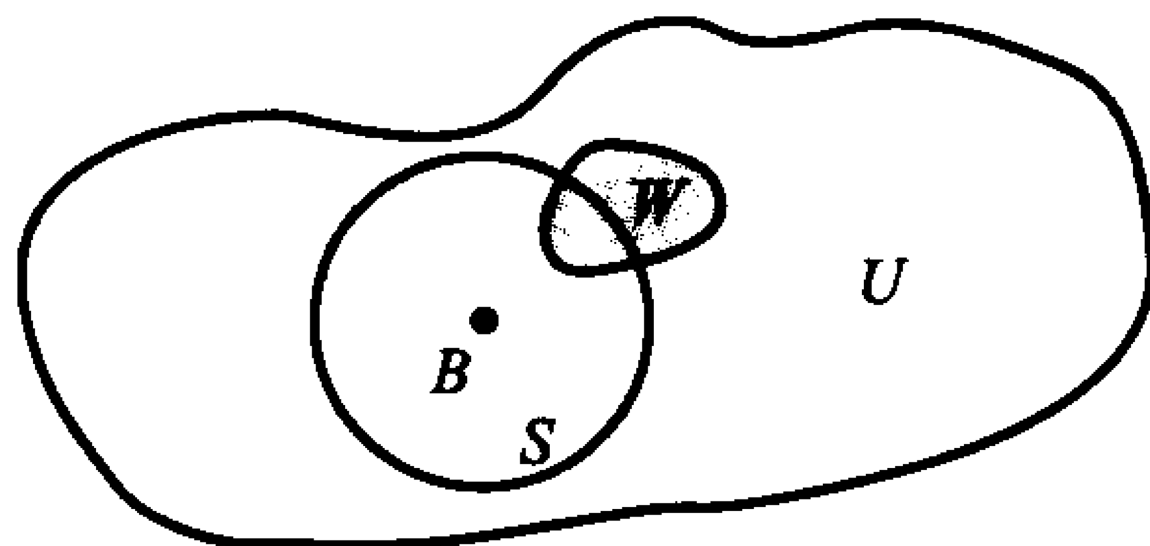


图 12.9

对所有 $x \in W$ 成立. 由连续性, 这表明 $w_B(y) - v(y) = w_B(x_0) - v(x_0)$ 对 $y \in W \bigcap S$ 成立. 如果 $W \bigcap S \neq \phi$, 我们可以取 y 为一个新的 x_0, 而问题则化到了前一种情形.

简而言之, 对定义在 U 上的每个次调和函数 w 我们联系了另一个次调和函数 w_B 使得 w_B 实际上在 B 的内部为调和而在 $U - B$ 中与 w 重合, 并在整个 U 上有 $w \leqslant w_B$. 从构造的方法上可以清楚看到, 如果 w_1 和 w_2 是定义于 U 上的两个次调和函数, 它们满足 $w_1 \leqslant w_2$, 则

$$w_{1B} \leqslant w_{2B}.$$

12.9 狄利克雷问题

设 U 为 $\mathbb{E}^n$ 中一个开子集使 $\overline{U}$ 为紧. 我们称 U 具有一个可切触边界是说对每个 $p \in \partial U$, 存在一个球 B 使得 $\overline{B} \bigcap \overline{U} = \{p\}$, 因而图 12.10(a) 代表了一个具可切触边界的区域, 而图 12.10(b) 和 12.10(c) 表现了在它们的边界上有不可切触的点

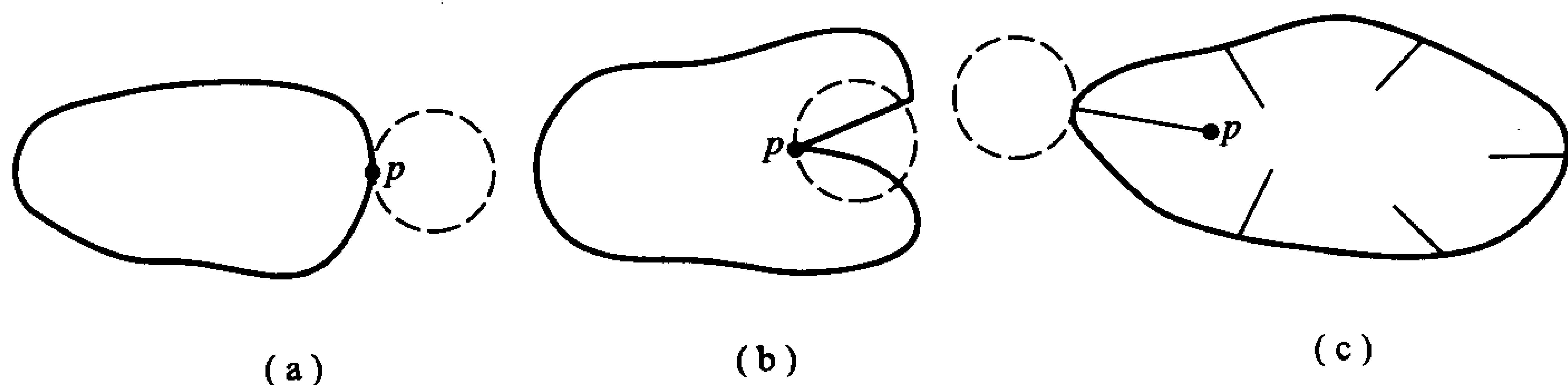


图 12.10

设 U 为 $\mathbb{E}^n$ 的开子集使 $\overline{U}$ 为紧, 并具有可切触边界. 设 f 为 ∂U 上的有界函数, 并假设对所有 $p \in \partial U$ 有 $|f(p)| < M$. 设 W_f 表示定义于 U 上并满足下面两个条件的所有函数的函数类. 这两个条件是:

(a) 每个 $w \in W_f$ 为次调和的; 以及

(b) 对每个 $p \in \partial U$,

$$\lim_{x \xrightarrow{x \in U} p} \sup w(x) \leqslant f(p).$$

我们可以重新叙述条件 (b) 如下: 对每个 $p \in \partial U$ 和每个 $\varepsilon > 0$, 存在一个 δ (依赖于 p, w, ε) 使得对 $\|x - p\| < \delta$ 有

$$w(x) \leqslant f(p) + \varepsilon.$$

注意, 因为恒等于 $-M$ 的常值函数显然属于 W_f, 故函数族 W_f 非空. 还需注意到, 条件 (b) 意味着当 $x \to \partial U$ 时 $\limsup w(x) < M$. 由于 $w \in W_f$ 是次调和的, 我们由极大值原理知道, 对所有 $w \in W_f$ 都有 $|w| < M$.

现在定义函数 u_f 为

$$u_f(x) = \operatorname*{lub}_{w \in W_f} w(x).$$

由前面的评注知道, 函数 u_f 是确有定义的, 并且事实上对所有 $x \in U$ 有 $|u(x)| < M$. 我们要证明, 如果 f 在 ∂U 上连续, 则函数 u_f 是对 U 的狄利克雷问题的解. 我们因而必须证明 u_f 是调和的并取边界值 f.

我们首先在对 f 没有任何连续性的限制下证明 u_f 是调和的. 要做到这点只要证明 u_f 在任意开球 B 中是调和的即可, 这里的 B 满足 $\overline{B} \subset U$. 设 x 为 B 中某点. 设 $w_1, w_2, \cdots, w_k, \cdots$, 为属于 W_f 的一个函数序列使得 $\lim w_i(x) = u_f(x)$.(由 u_f 的定义这样的序列是存在的.) 现在定义

$$\begin{aligned} w_1' &= w_1, \\ w_2' &= w_1 \vee w_2, \\ &\vdots \\ w_i' &= w_1 \vee \cdots \vee w_i. \end{aligned}$$

于是 $w_1' \leqslant w_2' \leqslant \cdots \leqslant w_i' \leqslant \cdots$ 是一个属于 W_f 的单调递增的函数序列, 满足 $\lim w_i'(x) = u_f(x)$. 现在将每个 w_i' 换为 w_{iB}'. 函数 w_{iB} 为次调和且在 ∂U 附近与 w_i' 重合, 故 w_{iB} 属于 W_f. 另外, 因为 $w \leqslant w_B$ 对任何次调和函数都成立, 我们得到 $\lim w_{iB}' = u_f(x)$. 在球 B 内部每个函数 w_{iB}' 为调和, 所以序列 $\{w_{iB}'\}$ 是 B 中调和函数的非降单调函数. 由于 W_f 中所有函数都以 M 为上界, 我们从命题 7.2 知, 序列 $\{w_{iB}'\}$ 在 B 的任何一个紧子集上一致收敛于 B 中调和的函数 u, 因而有 $u_f(x) = u(x)$. 另外, 由定义, 对任意 $y \in B$ 我们有 $u(y) \leqslant u_f(y)$. 我们要证明实际上对所有这些 y 成立 $u(y) = u_f(y)$. 大体上我们按在 y 的相同步骤进行, 即选取序列 $\{v_{iB}'\}$, 其中每个 v_{iB}' 在 B 中调和, 且属于 W_f, 并使得此序列为单调非降, 而且 $\lim v_{iB}'(y) = u_f(y)$. 如此一来, 我们得到另一个函数 v, 它在 B 中调和且在整个 B 上满足 $v \leqslant u_f$ 和 $v(y) = u_f(y)$. 最后, 设函数 s_i 的定义为

$$s_i = w_{iB}' \vee v_{iB}',$$

并考虑函数 s_{iB}. 一方面, 它们属于 W_f 并在 B 中调和. 另外, 我们有 $w_{iB}' \leqslant s_{iB}, v_{iB} \leqslant s_{iB}$. 最后, 序列 $\{s_{iB}\}$ 为非降, 因此收敛于 B 中的调和函数 s. 由构

造, 在整个 B 上有

$$u \leqslant s \qquad \text{和} \qquad v \leqslant s,$$

而 $u(x) = s(x) = u_f(x)$ 和 $u(y) = s(y) = u_f(y)$. 把极大值原理用于调和函数 $u-s$ 和 $v-s$, 得到 (因为在 x 和 y 分别都达到极大值 0)$s = u = v$.在整个 B 内成立. 于是对任意 $y \in B$ 有 $u_f(y) = u(y)$, 这表明 u_f 在 U 上为调和. 注意在证明 u_f 为调和中我们没有用到 U 的边界的任何性质, 或者关于 f 的任何连续性的假定.

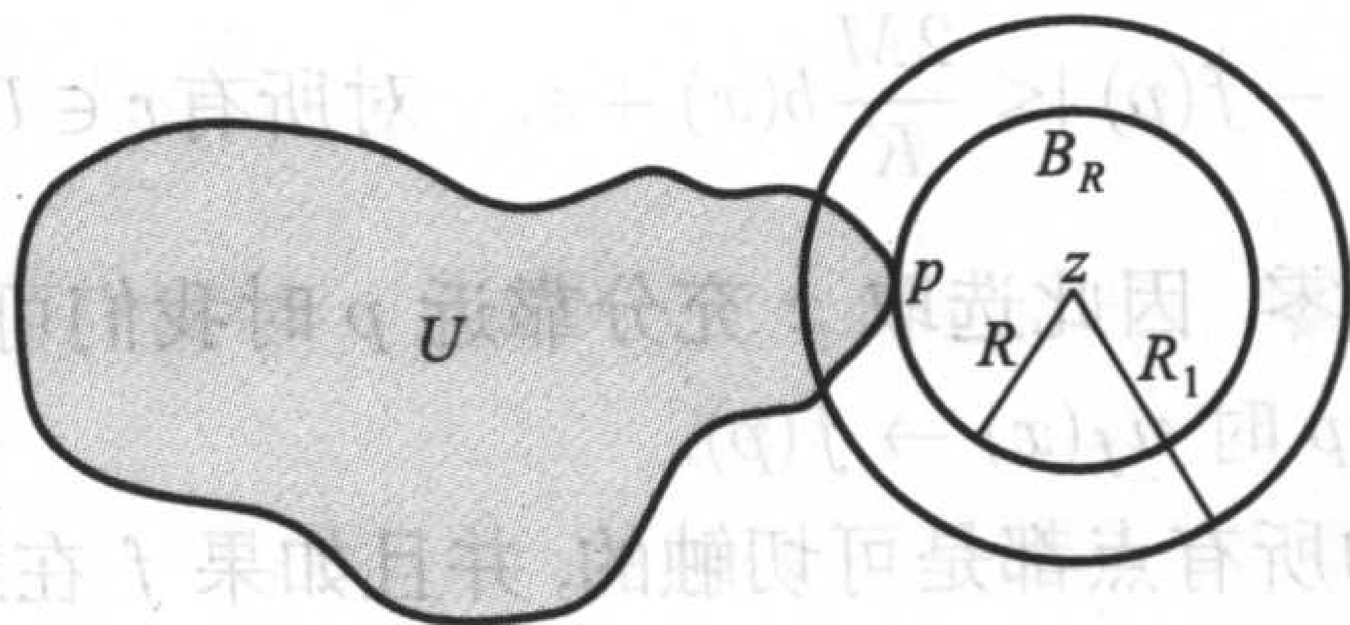


图 12.11

现在设 $p \in \partial U$ 为边界的一个切触点, 并假定 f 在 p 点连续. 我们要证明当 $x \to p$ 时 $u_f(x) \to f(p)$. 设 B_R 是半径为 R 中心在 z 的球, 它在 p 切触 $\overline{U}$, 即 $B_R \bigcap \overline{U} = \{p\}$(见图 12.11). 由于 f 在 p 连续, 故对任意 $\varepsilon > 0$ 可以找到一个 $R_1 > R$ 使得

$$|\, f(q) - f(p)\,| < \varepsilon, \quad \text{对所有} q \in \partial U \bigcap B_{R_1} \text{成立}. \tag{9.1}$$

定义函数 b 为 $b(x) = \|x-z\|^{2-n} - R^{2-n}$.[4] 注意 b 是定义于 $\mathbb{E}^n - \{z\}$, 并在其上为调和, 而且当 $\|x-z\| > R$ 时为负. 另外, 如果 $\|\, x-z\,\| \geqslant R_1$, 则 $b(x) \leqslant K < 0$. 特别地, 对所有 $q \in \partial U$, 我们有

$$f(q) < \frac{2M}{K} b(q) + \varepsilon + f(p). \tag{9.2}$$

事实上, 如果 $\|\, q-z\,\| < R_1$, 这正好是 (9.1) 式, 而如果 $\|\, q-z\,\| \geqslant R_1$, 则它由 $|f(q)| < M$ 得到. 由对次调和函数的极大值原理我们得出对任意 $w \in W_f$ 和任意 $x \in U$ 有 $w(x) < (2M/K)b(x) + \varepsilon + f(p)$. 因而我们有

$$u_f(x) < \frac{2M}{K} b(x) + f(p) + \varepsilon, \quad \text{对任意 } x \in U \text{ 成立}.$$

---

[4] 如果 $n = 2, b(x) = \ln(R/\|x-z\|)$.

另一方面, 按上面同样的理由得到

$$f(p)-\varepsilon-\frac{2M}{K}b(q)<f(q),\quad 对任意\ q\in\partial U\ 成立,$$

所以调和函数 $f(p)-\varepsilon-(2M/K)b$ 实际上属于 W_f. 特别, 我们有

$$f(p)-\varepsilon-\frac{2M}{K}b(x)\leqslant u_f(x),\quad 对所有\ x\in U\ 成立.$$

两个不等式放在一起, 我们得到

$$|u_f(x)-f(p)|\leqslant\frac{2M}{K}b(x)+\varepsilon,\quad 对所有 x\in U 成立. \tag{9.3}$$

函数 b 连续并在 p 为零. 因此选取 x 充分靠近 p 时我们可使得 (9.3) 式的右端小于 2ε. 于是当 $x\to p$ 时 $u_f(x)\to f(p)$.

特别, 如果 ∂U 的所有点都是可切触的, 并且如果 f 在边界上所有点是连续的, 我们就已经证明了:

定理 9.1 *设 U 为 $\mathbb{E}^n$ 中具紧闭包和可切触边界的开子集. 设 f 为定义在 ∂U 上的连续函数. 于是存在惟一的函数 u_f, 它在 $\overline{U}$ 上连续, 在 U 中调和, 并且在 ∂U 上与 f 重合.*

有一些关于定理 9.1 和它的证明的评注是适宜的.

(a) 我们用到 ∂U 是可切触的假定的只有一次, 即在我们构造函数 b 时, 我们要求它在 $\overline{U}$ 内为次调和, 在 p 点为零, 并且在 ∂U 上但在 p 的某个邻域外的所有点取的值小于某个负的常数. 一个更加仔细的分析指出, 只要我们能从外部用一个锥切触 p, 我们则总能够构造出这样一个函数. 因此我们能够在甚至像图 12.12 中那样的 p 点上解出狄利克雷问题.

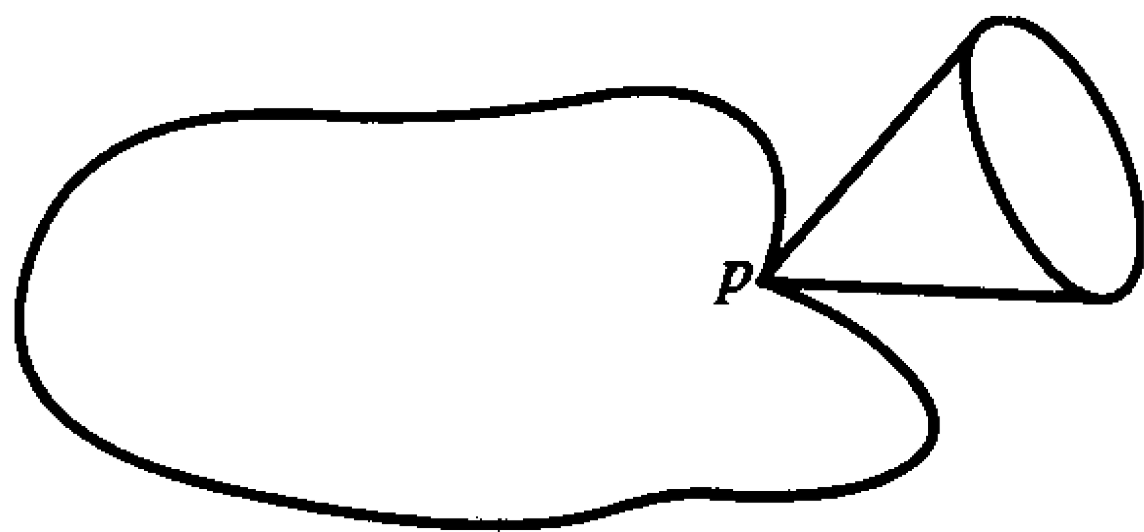


图 12.12

(b) 另一方面, 在边界上给出某种条件是必须的. 例如, 如果 ∂U 包含一个孤立点 p, 那么我们不能随意地指派给 p 一个值. 事实上, 如果 u 在 p 的一个邻

域 W 中连续, 且在 $W-\{p\}$ 为调和, 则泊松积分公式表明 u 的各种导数也要在 $W-\{p\}$ 中有界. 因此把中值定理的证明用于点 p 自身, 从而 $u(p)$ 被确定, 它便不能被随意指定. 更一般地, 一个更加精巧的论证指出, 对于边界包含了指向内部的尖钉或 (在 dim$\geqslant$ 3) 充分尖窄的尖点 (如图 12.13) 的区域, 狄利克雷问题不可能有解.

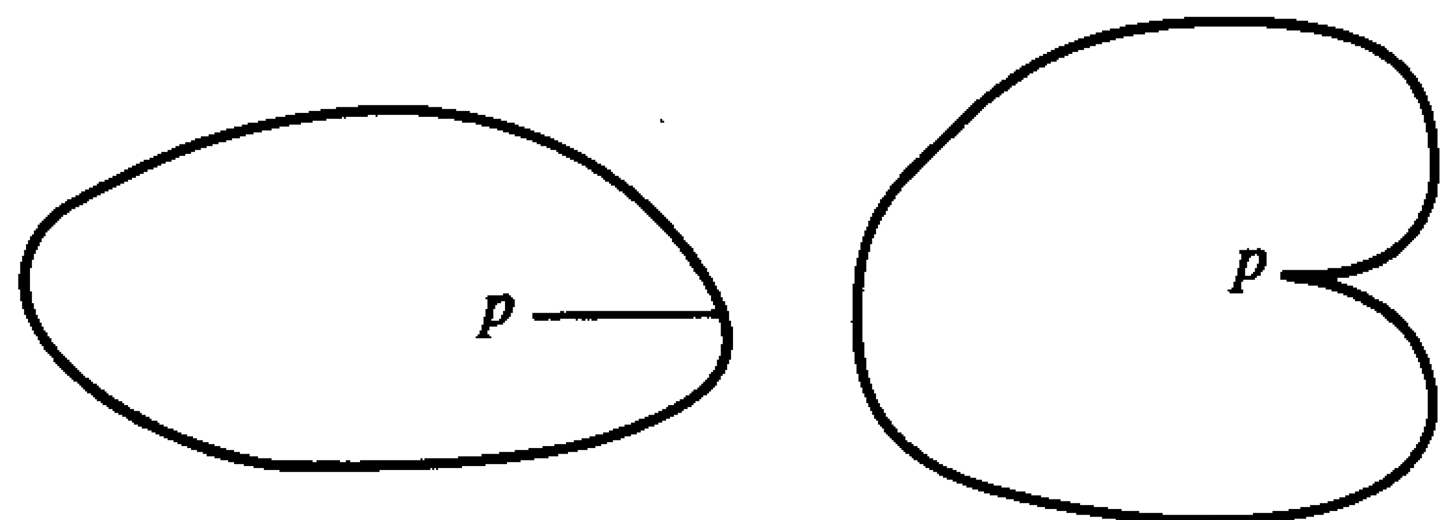


图 12.13

这是一个物理事实的数学类比, 这个物理事实说一个非常尖的导体不能充满电荷而会火花放电.(与静电学的关系将在第 12 节讨论.)

(c) 在应用第 4 节的结果方面即构造此区域的格林函数和各种等式时, 定理 9.1 仍然不十分够. 我们所需要知道的不仅仅是格林函数存在而且还要知道它有直到边界的连续导数. 这便需要进一步的论证和关于边界性质的更多的假设条件; 这将在下一节中讨论.

12.10 边界附近的行为

本节的目的是讨论在边界附近狄利克雷问题的解的行为. 事实上我们要证明:

定理 10.1 *设 U 为具有正则边界的区域, 且在 $\mathbb{E}^n$ 中有紧闭包. 设 f 为定义在 ∂U 的连续可微函数, 而 u 是具边界值 f 的狄利克雷问题的解. 那么, 偏导数 $\partial u/\partial x^i$ 可以扩张到 $\overline{U}$ 上的连续函数. 另外, 如果 $p\in\partial U, \xi=\prec\xi^1,\cdots,\xi^n\succ$ 在 p 点切于边界, 则*

$$\prec\xi,df\succ=\prec\xi,du\succ=\sum(\partial u/\partial x^i)\xi^i.$$

下面的证明是由阿尔福斯教授向我们提出来的, 经他好意的允许重新用于此处.

证明 设 p 为 ∂U 的某个点, 经欧几里得运动, 我们重新安排 ∂U 在 p 的切平面为平面 $x^n = 0$ (见图 12.14). 于是 ∂U 的在 p 点附近的点是所有形如 $(x^1, \cdots, x^{n-1}, \varphi(x^1, \cdots, x^{x-1}))$ 的点, 其中 φ 是在 $\mathbb{R}^{n-1}$ 的原点附近定义的函数, 它在原点连同它的所有一阶导数场为零. 特别地, 存在常数 $C > 1$ 使得在 $\mathbb{R}^{n-1}$ 的原点的某个邻域中有

$$|\varphi(x^1, \cdots, x^{n-1})| \leqslant C\{(x^1)^2 + \cdots + (x^{n-1})^2\}.$$

我们因此可选取一个充分小的 $R < 1/2C$ 使得半径为 R 中心在 $(0, \cdots, 0, R)$ 的 (开) 球完全在 U 内, 而半径为 R 中心在 $(0, \cdots, 0, -R)$ 的球在 $\overline{U}$ 的补集内 (见图 12.15).

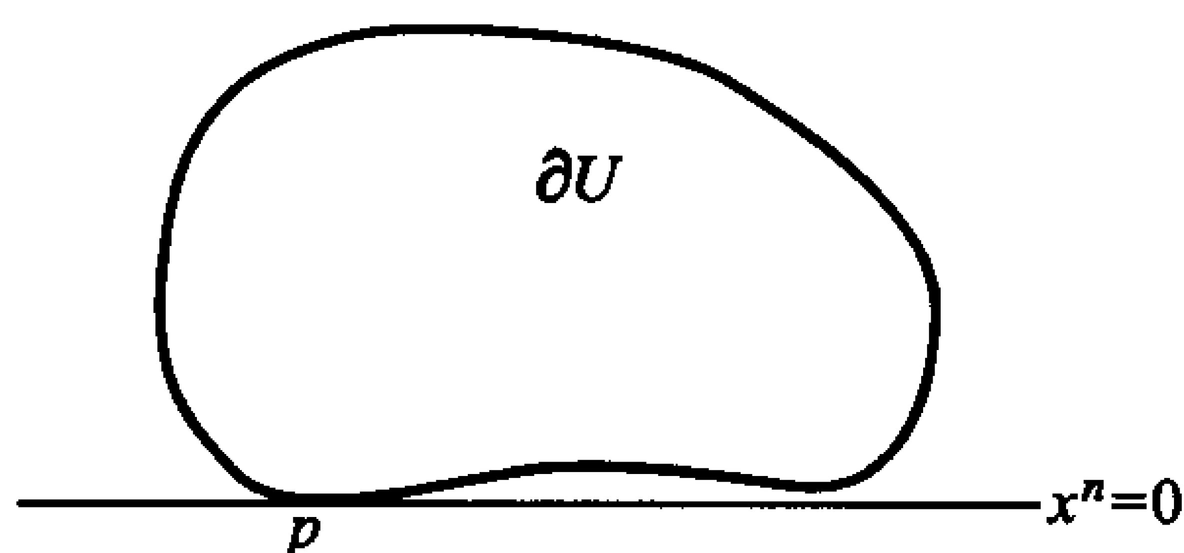


图 12.14

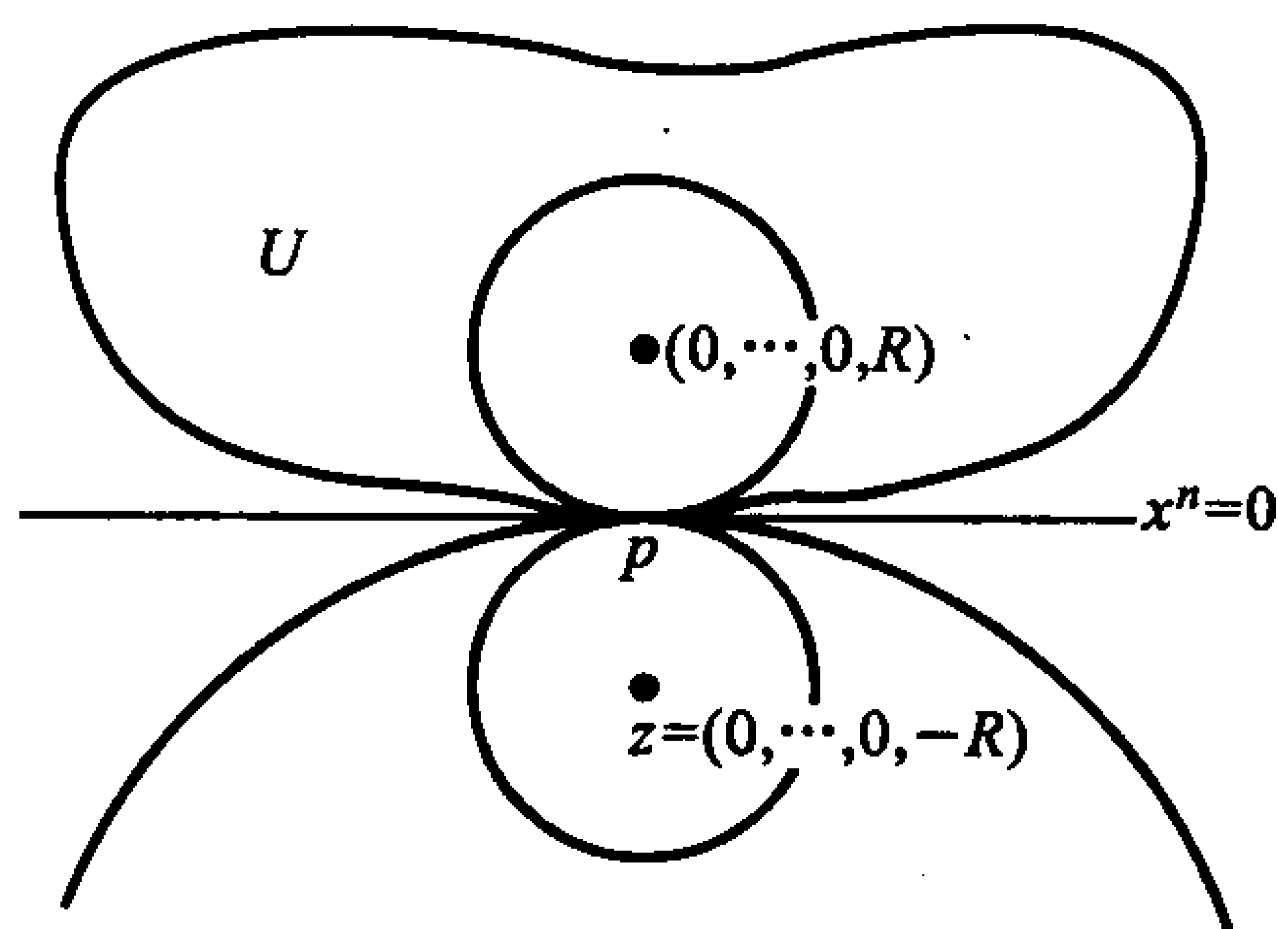


图 12.15

设 $z = (0, \cdots, 0, -R)$. 于是对所有充分靠近 p 的 $q \in \partial U$ 我们有

$$\begin{aligned}||q - z||^2 &= (x^1)^2 + \cdots + (x^{n-1})^2 + R^2 - 2R\varphi(q) + \varphi^2(q) \\ &\geqslant (1 - 2RC)\{(x^1)^2 + \cdots + (x^{n-1})^2\} + R^2 + \varphi^2(q),\end{aligned}$$

从而

$$||q - z|| - R \geqslant k\{(x^1)^2 + \cdots + (x^{n-1})^2\},$$

其中 k 是某个正常数. 设 $\varphi(r) = r^{n-2}$(如果 $n = 2$, 则为 $\ln r$). 于是 $\varphi'(R) \neq 0$ 从而 $|\varphi(||q - z||) - \varphi(R)| > C|||q - z|| - R|$, 其中 $C > 0$ 为某个适当的常数. 特别, 我们有不等式

$$\left|\frac{1}{||q - z||^{n-2}} - \frac{1}{R^{n-2}}\right| \geqslant K\{(x^1)^2 + \cdots + (x^{n-1})^2\} \tag{10.1}$$

或者当 $n=2$ 时

$$\left|\ln\frac{\|q-z\|}{R}\right| \geqslant K(x^1)^2,$$

其中 K 为某个正常数.

现在让我们把函数 f 换为 $f-[f(p)+\sum_1^{n-1}(\partial f/\partial x^i)(p)x^i]$. 于是 $u-[f(p)+\sum_1^{n-1}(\partial f/\partial x^i)(p)x^i]$ 是相应的狄利克雷问题的解. 如此一来, 我们可假定 (相应地改变我们的记号) f 连同它的一阶导数在 p 为零. 由于假定了 f 有连续的一阶导数, 我们便可以应用泰勒定理得出结论说, 存在 $c>0$ 使得对所有充分靠近 p 的 ∂U 上的 q 我们有

$$|f(q)| \leqslant c\{(x^1)^2+\cdots+(x^{n-1})^2\}. \tag{10.2}$$

像在上一节那样, 设

$$b(x)=\|x-z\|^{2-n}-R^{2-n}$$

$$\left(\text{或当 } n=2 \text{ 时}, b(x)=\ln\frac{R}{\|q-z\|}\right).$$

如果比较 (10.2) 和 (10.1) 式, 我们看出对所有充分靠近 p 的 $q\in\partial U$ 我们有

$$|f(q)| \leqslant \frac{c}{K}|b(q)|.$$

另一方面, 因为 b 在 p 的某个邻域外为严格的负值 (在 ∂U 上), 并且因为 f 在 ∂U 上有界, 我们可以 (对所有靠近 p 的 q 应用上面的不等式) 找到一个常数 A 使得

$$|f(q)| \leqslant A|b(q)|, \qquad \text{对所有} q\in\partial U \text{成立}. \tag{10.3}$$

因为函数 b 在 U 中为调和的, 将极大值原理用于 $u-Ab$ 和 $Ab-u$ 使我们得到

$$|u(x)| \leqslant A|b(x)|, \qquad \text{对所有} x\in\overline{U} \text{成立}. \tag{10.4}$$

现在函数 b 是从 z 到 x 的距离的可微函数, 当距离为 R 时它为零. 特别地, 存在常数 $a>0$ 使得

$$|b(x)| \leqslant ad(x),$$

其中 $d(x)$ 代表从 x 到中心在 z, 半径为 R 的球面的距离. 设 B 为半径为 R 中心在 $-z=(0,\cdots,0,R)$ 的球, 故 B 在 U 中, 设 $S=\partial B$. 注意, S 在 p 点切于中

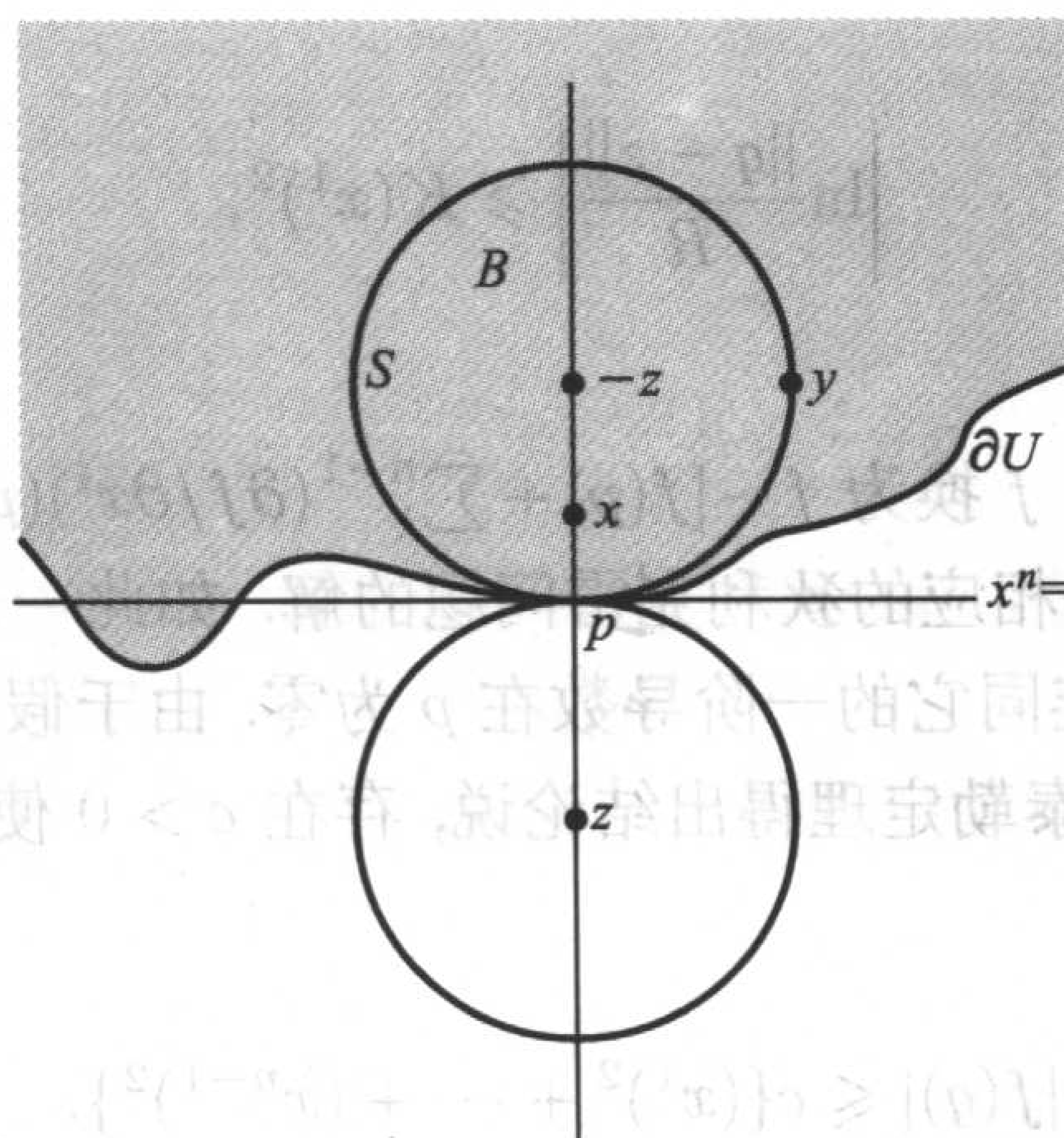


图 12.16

心在 z 的球 (见图 12.16). 因此对 S 上的点 y 有 $d(y) \leqslant c_1\|y-p\|^2$. 如果运用前面的不等式把这个代入 (10.4) 式, 我们得到

$$|u(y)| \leqslant L\|y-p\|^2, \qquad \text{对所有} y \in S \text{成立}, \tag{10.5}$$

其中 L 为某个常数. 将泊松积分公式用于球 B 和函数 u 并作关于 x^i 的微分, 我们有

$$\frac{\partial u}{\partial x^i} = \frac{-2(x^i+z^i)}{A_{n-1}R}\int \frac{u(y)}{\|x-y\|^n}dS + \frac{R^2-\|x+z\|^2}{A_{n-1}R}\int \frac{u(y)(x^i-y^i)}{\|x-y\|^{n+2}}dS. \tag{10.6}$$

现在设 x 在过 p 的边界的法线上, 故 $x=(0,\cdots,0,x^n)$. 如果 $|x^n|<R$, 则

$$\|y-p\| \leqslant \|y-x\| + |x^n| \leqslant 2\|y-x\|.$$

因此

$$\left|\int \frac{u(y)}{\|x-y\|^n}dS\right| \leqslant 2^n \int \frac{|u(y)|}{\|p-y\|^n}dS \leqslant 2^n D \int \|y-p\|^{2-n}dS.$$

因为此球面是 $n-1$ 维的, 最后面的一个积分绝对收敛, 从而我们看到在 (10.6) 式中第一个积分一致绝对收敛.(注意, 当 $i<n$ 时包含这个积分的项为零.) 我们现在证明 (10.6) 式的第二项为 x^n 趋向零时趋向于零. 事实上, $R^2-\|x+z\|^2=$

$x^n(2R - x^n)$, 故我们可将 (10.6) 式的第二项写为

$$\frac{(2R - x^n)(x^n)^{1/2}}{A_{n-1}R} \int \frac{(x^n)^{1/2} u(y)(x^i - y^i)}{\|x - y\|^{n+2}} dS.$$

这一回因为 $x^n \leqslant \|x - y\|$ 对所有 $y \in S$ 成立, 我们则可断定此被积项的绝对值小于

$$\frac{|u(y)||x^i - y^i|}{\|y - x\|^{n+3/2}} \leqslant \frac{|u(y)|}{\|y - x\|^{n+1/2}} \leqslant \frac{|u(y)|2^{n+1/2}}{\|y - p\|^{n+1/2}}$$
$$\leqslant D 2^{n+1/2} \|y - p\|^{3/2-n},$$

它还是绝对可积的. 因此出现在最后那个表达式中的积分对所有 $x^n \geqslant 0$ 的值是一致有界的. 由于 $(x^n)^{1/2} \to 0$, 我们得到了 (10.6) 式中第二项趋向于 0 的结论.

如果我们现在对所得到的结果不用特定的坐标选取来重新措词的话, 我们看到我们已经证明了下面的结果: 设 p 为 ∂U 中任意点, x 为 ∂U 的通过 p 的法线上的一点. 如果 ξ 是任意一个平行于在 p 点的切向量的向量, 则当沿着此法线 $x \to p$ 时 $\langle \xi, du \rangle \to \langle \xi, df \rangle$. 如果 $\partial/\partial n$ 为法方向的单位向量 (指向 U 内), 则

$$\left\langle \frac{\partial}{\partial n}, du \right\rangle \longrightarrow \frac{-2}{A_{n-1}} \int \frac{u(y) - u(p)}{\|y - p\|^n} ds, \tag{10.7}$$

这里的积分是取在半径为 R, 在 p 点切于 ∂U 并且位于 $\overline{U}$ 内的球面上.

迄今我们已经证明了仅当沿着法方向 $x \to p$ 时的收敛性. 如果我们返回查验这个论证, 我们看到半径 R 和 (10.5) 式中的常数 D 可以对所有 $p \in \partial U$ 进行一致地选取. (事实上, 因为 ∂U 为紧, 我们可以用有限个 (iii) 型的邻域覆盖 ∂U, 使得围绕每个 $p \in \partial U$ 的半径为 R 的球位于这些邻域中的一个, 这时的 R 充分小. 在这种情形下可清楚看出 R 的选取 (如果必要仍旧很小) 仅仅依赖于从欧氏坐标到已校正分图表的分图表变换的二阶导数. D 的选取也仅仅依赖于包含在 f 的泰勒展开中的常数和 R. 因此 R 和 D 的选取可以做成是一致的.) 由于按假定条件, f 的偏导数连续, 而且 (10.7) 式右端对 p 连续, 我们最后得出, 当按任意方向趋向边界时 u 的这些偏导数的极限存在, 从而事实上我们证明了定理 10.1. □

现在我们对任何具正则边界的区域, 用狄利克雷问题的解, 按第 4 节那样构造一个格林函数. 定理 10.1 告诉我们, 它在闭集 $\overline{U}$ 内是可微的, 这个意义是说偏导数直到边界都是连续的. 如果我们需要推导出第 4 节的各种公式, 我

们可应用极限论证法来做: 只要简单地把 U 换成更小的区域 U^a 使 $a \to 0$ 时 $U^a \to U, \partial U^a \to \partial U$.

(实际上, 一个更仔细更精妙的论证指出, 如果 f 具有两个连续导数, 则我们作为在边界上的极限得到的 $\partial u/\partial x^i$ 事实上是连续可微的. 因为 $\partial u/\partial x^i$ 是具这些边界值的狄利克雷问题的解, 我们得出的结论是 u 的二阶导数可以扩张使其在 $\overline{U}$ 上连续. 按此办法可以证明, 如果 f 为 C^∞, 则 u 的所有导数可以扩张到在 $\overline{U}$ 上连续. 特别地, 格林函数 $G(x,\cdot)$ 的所有导数可以扩张到边界使对任一 $x \in U$ 为连续函数.)

12.11 狄利克雷原理

狄利克雷问题的解有一个有趣而又有用的特征刻画, 即它是在已知边界值下使狄利克雷积分 $D[u,v] = \int \sum |\partial u/\partial x^i)^2$ 极小的问题的解. 更准确地说,

定理 11.1 (狄利克雷原理) 设 U 为具紧闭包和正则边界的区域, f 为 ∂U 上连续可微的函数. 设 u 为以 f 为边界值的狄利克雷问题的解, 而 w 是任意一个在 U 中可微的函数, 它在边界上取 f 的值, 而且它的导数可扩张为 $\overline{U}$ 上的连续函数. 于是

$$D[u,u] \leqslant D[w,w], \tag{11.1}$$

并且等式成立当且仅当 $u = w$.

证明 记 $w = u + v$, 那么 v 在 U 连续可微, 在 $\overline{U}$ 上连续, 而在 ∂U 上为零. 于是由 D 对其变量的双线性性和对称性.

$$D[w,w] = D[u+v,u+v] = D[u,u] + D[v,v] + 2D[u,v]. \tag{11.2}$$

现暂且假定 u 具有两个连续导数, v 具有在 $\overline{U}$ 的一个邻域中的一个连续导数. 于是

$$D[u,v] = D_U[u,v] = \int_U dv * du = \int_{\partial U} v * du - \int v\Delta u = 0. \tag{11.3}$$

而 $D[v,v] = \int \sum (\partial v/\partial x^i)^2 \geqslant 0$, 其为 0 当且仅当 v 的所有导数为零, 故 v 在 U 的每个分支上为常数. 由于 v 在 ∂U 上为 0, 故 $D[v,v] = 0$ 当且仅当 $v = 0$.

如果 u 和 v 在 ∂U 的邻域中不一定是可微的, 我们记

$$D_U[u,v] = \lim D_{U^a}[u,v],$$

其中 U^a 是当 $a \to 0$ 时趋向于 U 的一系列区域, 同时 $\partial U^a \to \partial U$, 则可建立 $D[u,v]=0$. 这时应用前面的论证到每一个 U^a 上便得出 $D[u,v]=0$, 和 $D[v,v]=0$ 当且仅当 $v=0$ 的结论. 定理得证. □

12.12 物理应用

在经典物理的许多理论中, 拉普拉斯方程和它的解的研究起了重要作用. 这主要由于拉普拉斯算子和欧几里得几何之间的紧密联系. 由于这不是详细阐述物理应用的场合, 我们推荐读者参考任一本物理教材以获得详细的内容.(例如可以看 Feymman, Leighton 和 Sands 所著的《The Feynman Lectures on Physics》, Addison-Wesley, Reading, Mass., 1964, Vol II, 特别是第十二章.) 在这节我们将简短地提及一些相关的物理.

在静电学中假定了电场 E 满足方程

$$d*E=\rho, dE=0$$

其中 ρ 是电荷密度, 而且把向量场 E 用 $\mathbb{E}^3$ 上的欧氏度量等同于一个线性微分形式. 这里的 ρ 被暂且假定为一个光滑的密度. (另外考虑 "曲面分布" 和 "点分布" 这些极限情形也颇方便,) 由上面第二个方程我们可知 E 能写成

$$E=-d\rho,$$

其中 φ 是个光滑函数, 称为电场的电势. 于是第一个方程表示了

$$\Delta\varphi=d*d\varphi=-\rho$$

如果已知 ρ, 则 φ 由公式

$$\varphi(y)=\frac{1}{4\pi}\int\frac{\rho(x)}{||y-x||}dx$$

决定 (因 $A_2=4\pi$). 作为极限情形, 对于以密度 σdA 沿曲面 S 的电荷分布我们得到公式

$$\varphi(y)=\frac{1}{4\pi}\int_S\frac{\sigma(x)}{||y-x||}dA,$$

其中 dA 是此曲面的面积密度, 而

$$\varphi(y)=\frac{1}{4\pi}\int_C\frac{l(y)}{||y-x||}dS,$$

其中 ldS 是沿曲线 C 的线分布. 对在 x 的一个点电荷, 我们有

$$\varphi(y) = \frac{e}{4\pi\|y-x\|},$$

其中 e 是电量.

沿着一个导体可以没有电场 (因为如果有电场就会导致电荷运动, 从而与我们的静电场假定矛盾). 因此

$$d\varphi = 0, \qquad \text{从而 } \varphi \text{ 为常数.}$$

在大多数的问题中电荷分布的产生不是事先就了解的, 然而却必须确定它. 例如, 假定在点 x 处有一个单位电荷, 其中 x 位于一个腔体内部, 此腔体边界是个接地的导体, 即具零电势 (见图 12.17). 我们要求出腔体内部的电场. 我们也希望确定在边界曲面上所诱导的电荷分布. 因为在此腔体内部除了 x 之外没有任何地方有电荷分布, 故电势 φ 必定除 x 外处处为调和, 并且事实上与

$$\frac{1}{4\pi\|\cdot - x\|}$$

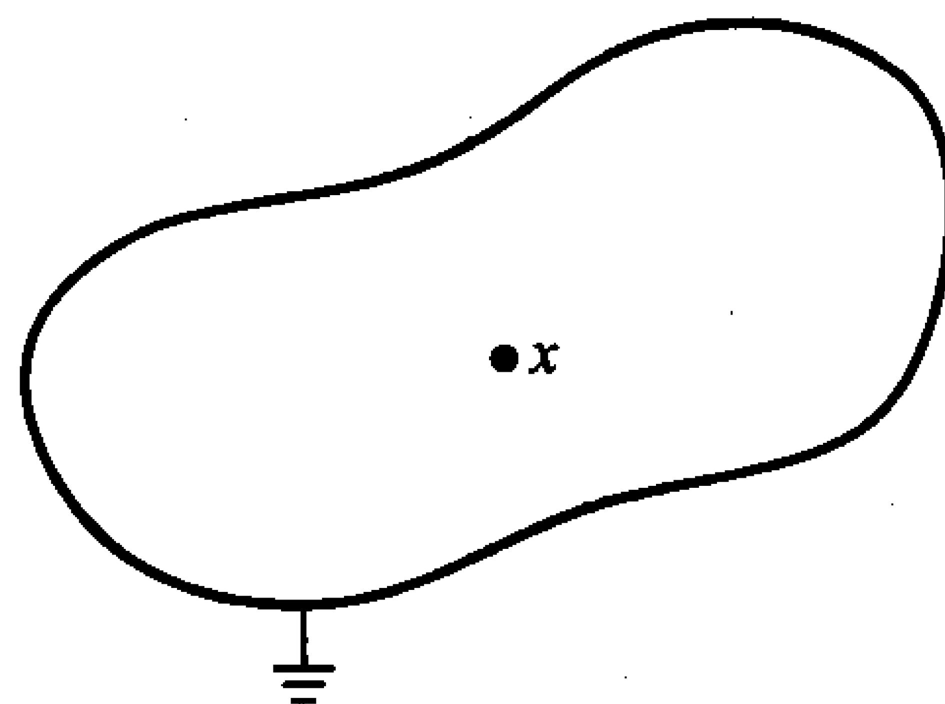


图 12.17

相差一个调和函数. 因为在边界上 $\varphi = 0$, 所以我们知道所要的电势 φ 完全是此区域的格林函数. 曲面密度从而可由 $d\varphi$ 沿此边界决定.(这是为什么了解狄利克雷问题在边界可微具有重要性的另一个理由.)

我们更经常地是对传导曲面外部的电场感兴趣. 严格地说, 我们所阐述的狄利克雷问题在这里用不上, 因为我们只考虑了有界区域. 可以用两种办法来处理这个问题. 首先, 可以考虑一个非常大的传导球面 (电势为 0) 内部所包括的一切东西; 从而把问题化为狄利克雷问题. 然后让半径趋向无穷大从而作为极限得到所想要的电势. 其次, 可以修改我们在第 3 节到第 10 节的论证使它们包含无界区域的情形, 但是必须限制所有我们的函数当 c, r 趋向无穷时为零. 详细情形留给读者.

在散射理论中假定了具某种特性的物质按照一个向量场 X 流动. 如果此物质的密度为 ρdV, 则在单位时间通过曲面 S 的一个定向片流动的物质总量由

$$\int_S \rho(X, \boldsymbol{n})dA = \int_S *\rho X$$

给出, 其中 n 是此曲面的单位法向量, dA 是面积元, 并且我们通过欧几里得度量把右端的 ρX 看作为线性微分形式. 因此流出任一区域的物质的净流量为 $d*\rho X=\text{div}\prec X,\rho\succ$. 现在, "物质" 可以从此区域内以速率 sdV 产生 (其中 s 是描述此区域中源头的生产速率的函数). 因此总的密度变化由

$$\frac{\partial\rho}{\partial t}dV=\text{div}\prec X,\rho\succ -sdV$$

给出, 如果像我们要假定的那样, 它是稳态的, 即密度不变, 我们则得到方程

$$\text{div}\prec X,\rho\succ = sdV.$$

另一方面, 常常遇到的流是由某个函数的梯度给出, 即

$$\rho_{\text{id}}X=dN,$$

N 为某个函数. 把这些等式结合起来我们得到

$$\Delta N=s.$$

在热力学理论中 N 是温度, $\rho_{\text{id}}X$ 是热流的速率, s 是热源的密度. 方程 $\rho_{\text{id}}X=dN$ 说的是热流的速率与温度的梯度成比例. 在质点散射理论中假定了物质流的速率与质点密度的变化 (即梯度) 成比例.(这就是说质点是从高密度区域流向低密度的区域.) 于是 N 代表了质点的密度而 s 代表了质点产生的速率.

例如, 假设区域 D 的边界保持在某个固定温度 f, 其中的 f 是 ∂D 上的一个已知函数, 并且在 D 内部没有热源. 于是在 D 中的温度分布是由满足微分方程 $\Delta N=0$ 的函数 N 给出, 边界条件是在 ∂D 上 $N=f$. 换句话说, 温度由狄利克雷问题的解决定.

在*连续, 不可压缩, 无旋的流体流*理论中假定了一个代表了液体流动的已知向量场 X. 由不可压缩性得到 $d*X=0$. 假定了无环流, 即 $dX=0$. 因此 $X=d\varphi$, 其中 φ 为拉普拉斯方程的一个解函数. 在这种情形下产生的自然的边界值问题不同于我们一直在讨论的那些, 所以我们直接让读者去参看关于流体力学的标准著作.

附加习题

1. 设 V_1 和 V_2 为具内积的向量空间. 线性映射 $l:V_1\to V_2$ 被称作保形 (或相似) 是说

$$(l_v,l_w)=c(v,w),$$

其中 c 是个正常数.(注意, 当 $c=1$ 时 l 为一个等距.)

设 M_1 和 M_2 为具黎曼度量的流形. 称可微 (C^∞) 映射 φ 为保形是说 φ_{*x} 在每点 $x \in M_1$ 是保形的. 因此, 说 φ 为保形意味着对任意 $x \in M_1$,

$$(\varphi_{*x}\xi_1, \varphi_{*x}\xi_2) = c(x)(\xi_1, \xi_2),$$

其中 $\xi_1, \xi_2 \in T_x(M_1)$.(注意这次的 C 依赖于 x.)

(a) 设 $\varphi: U \to \mathbb{R}^2$ 为保形映射, 其中 $U \subset \mathbb{R}^2$, 并且在 U 和 $\mathbb{R}^2$ 上我们都用欧氏度量. 假设 φ 由 $\varphi(x,y)=(u,v)$ 给出, 其中 $u=u(x,y), v=v(x,y)$, 而 (x,y) 和 (u,v) 都为此平面上的直角坐标. 证明方程

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

成立, 或者方程

$$\frac{\partial u}{\partial x} = -\frac{\partial v}{\partial y}, \frac{\partial u}{\partial y} = \frac{\partial v}{\partial x}$$

成立.

(b) 得出结论说, 如果 u, v 如同 (a) 中那样, 则它们都是调和函数.

2. (a) 设 u 是定义在整个 $\mathbb{R}^n$ 上的调和函数. 证明如果 u 有界则它必为常数.

(b) 运用 (a) 和习题 1, 证明没有 $\mathbb{R}^2$ 到一个有界子集 $U \subset \mathbb{R}^2$ 的保形映射.

12.13 问题汇编: 留数计算

在一个流形 M 上我们可以考虑复数值的光滑函数, 微分形式, 等等. 例如, 我们说函数

$$f = u + iv$$

是个复值 C^∞ 函数是指每个实函数 u 和 v 为实的 C^∞ 函数. 相似地, 我们考虑复值线性微分形式

$$\omega = \sigma + i\tau,$$

其中 $\sigma, \tau \in \wedge^1(M)$ 为 (实值) 线性微分形式. 于是我们可以进行通常的运算, 但要记住 $i^2=-1$. 因此

$$f\omega = (u\sigma - v\tau) + i(v\sigma + u\tau).$$

相似地, 如果 $\omega_1 = \sigma_1 + i\tau_1, \omega_2 = \sigma_2 + i\tau_2$, 则

$$\omega_1 \wedge \omega_2 = (\sigma_1 \wedge \sigma_2 - \tau_1 \wedge \tau_2) + i(\sigma_1 \wedge \tau_2 + \tau_1 \wedge \sigma_2)$$

为复值 2 - 形式.

我们类似地还有算子d, 它由

$$df = du + idv, \quad dw = du + id\tau,$$

等等给出.

如果 X_1 和 X_2 为向量场, 我们定义 "复向量场" $X_1 + iX_2$ 为在复值函数上的微分算子, 它由

$$(X_1 + iX_2)f = X_1 f + iX_2 f = (X_1 u - X_2 v) + i(X_1 v + X_2 u)$$

给出.

从现在起, 设 M 是 $\mathbb{R}^2$ 的一个开子集, 其直角坐标为 (x, y). 令 $z = x + iy$, 故 $dz = dx + idy$. 我们设

$$\bar{z} = x - iy, \quad d\bar{z} = dz - idy,$$

于是, 定义复向量场 $\partial/\partial z$ 和 $\partial/\partial \bar{z}$ 为

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \quad 和 \quad \frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

然后再设, 对任意复函数 f,

$$\frac{\partial f}{\partial z} = \left(\frac{\partial}{\partial z}\right) f \quad 和 \quad \frac{\partial f}{\partial \bar{z}} = \left(\frac{\partial}{\partial \bar{z}}\right) f$$

习题

13.1 证明对任意复值光滑函数 f 我们有

$$df = \frac{\partial f}{\partial z}dz + \frac{\partial f}{\partial \bar{z}}d\bar{z}.$$

13.2 证明对 $\partial/\partial z$ 和 $\partial/\partial \bar{z}$ 的莱布尼茨规则成立: 即

$$\frac{\partial (fg)}{\partial z} = f\left(\frac{\partial g}{\partial z}\right) + g\left(\frac{\partial f}{\partial z}\right), \frac{\partial (fg)}{\partial \bar{z}} = f\left(\frac{\partial g}{\partial \bar{z}}\right) + g\left(\frac{\partial f}{\partial \bar{z}}\right),$$

称一个函数 f 是全纯的(或复解析的) 是说 $\partial f/\partial \bar{z} = 0$. 对于全纯函数 f, 我们记

$$f' = \frac{\partial f}{\partial z}.$$

13.3 证明 dz 和 $d\bar{z}$ 在下面的意义下是独立的, 即

如果 $adz+bd\bar{z}=0$ 则, $a=0, b=0$.

13.4 证明结论: f 为全纯当且仅当对某个 h 有 $df=hdz$.

13.5 证明 $d(z^n)=nz^{n-1}dz$(如果 $n<0$ 则在 $\mathbb{R}^2$-$\{0\}$中), 其中 n 为所有的整数; 因而 z^n 为全纯.

13.6 证明由 $p(z)=\sum_0^n a_k z^k$ 给出的所有多项式 p 是全纯的, 且

$$p'(z)=\sum_1^n ka_k z^{k-1}.$$

13.7 定义函数 e^z 为 $e^z=e^x(\cos y+i\sin y)$. 证明 $de^z=e^z dz$.

13.8 证明两个全纯函数的乘积为全纯.

13.9 设 $f=u+iv$ 为全纯函数, 考虑 $(x,y)\mapsto(u,v)$ 为 $M\subset\mathbb{R}^2$ 到 $\mathbb{R}^2$ 的一个映射. 证明这个映射是保形的, 并且它在点 (x,y) 的雅可比行列式为 $\|f'(z)\|^2$, 这里的 $z=x+iy$.

13.10 设 g 为一个全纯函数, 它定义于 M 在对应于 f 给出的 $M\subset\mathbb{R}^2$ 的映射的像上. 定义 $g\circ f$ 为 $g\circ f(x,y)=g(u(x,y),v(x,y))$[为简明起见, 我们将其记为 $g\circ f(z)=g(f(z))$]. 证明 $g\circ f$ 全纯, 且

$$(g\circ f)'(z)=g'(f(z))f'(z).$$

13.11 设 U 为平面上具几乎正则边界的区域, 由斯托克斯定理, 对于任何复值线性微分形式 ν, 我们有

$$\int_{\partial U}\nu=\int_U d\nu.$$

(像平常那样, 可以解释为两端的实部和虚部的等式.) 证明

$$\int_{\partial U}gdz=\int_U dg\wedge dz=\int_U\frac{\partial g}{\partial\bar{z}}d\bar{z}\wedge dz.$$

13.12 证明 $d\bar{z}\wedge dz=2idx\wedge dy$, 故

$$\int_{\partial U}gdz=2i\int_U\frac{\partial g}{\partial z}dx\wedge dy=2i\int_U\frac{\partial g}{\partial t}.$$

13.13 证明如果 f 为定义于 $\overline{U}$ 的一个邻域上的光滑函数, 并且如果 f 在 U 中为全纯, 则 $\int_{\partial U}fdz=0$.

13.14 设 $\xi=\zeta+i\eta,(\zeta,\eta)\in U, g$ 为定义在 $\overline{U}$ 的某邻域上的光滑函数 (见图 12.18). 证明

$$g(\xi)=\frac{1}{2\pi i}\left[\int_{\partial U}\frac{g(z)}{z-\xi}dz+\int_U\frac{\partial g/\partial\bar{z}}{z-\xi}dz\wedge d\bar{z}\right].$$

[提示: 将习题 13.11 用于函数 $g(z)/z-\xi$ 和区域 $U-B_\varepsilon$, 其中 B_ε 是围绕 ξ 的半径为 ε 的一个圆盘. 然后让 $\varepsilon\to 0$.]

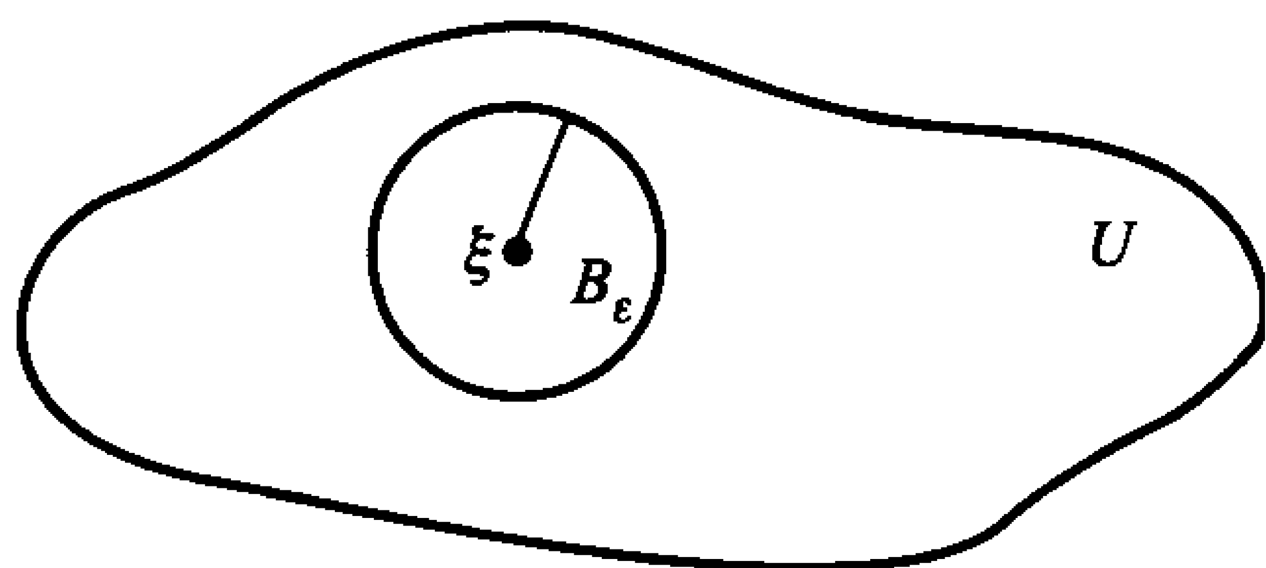


图 12.18

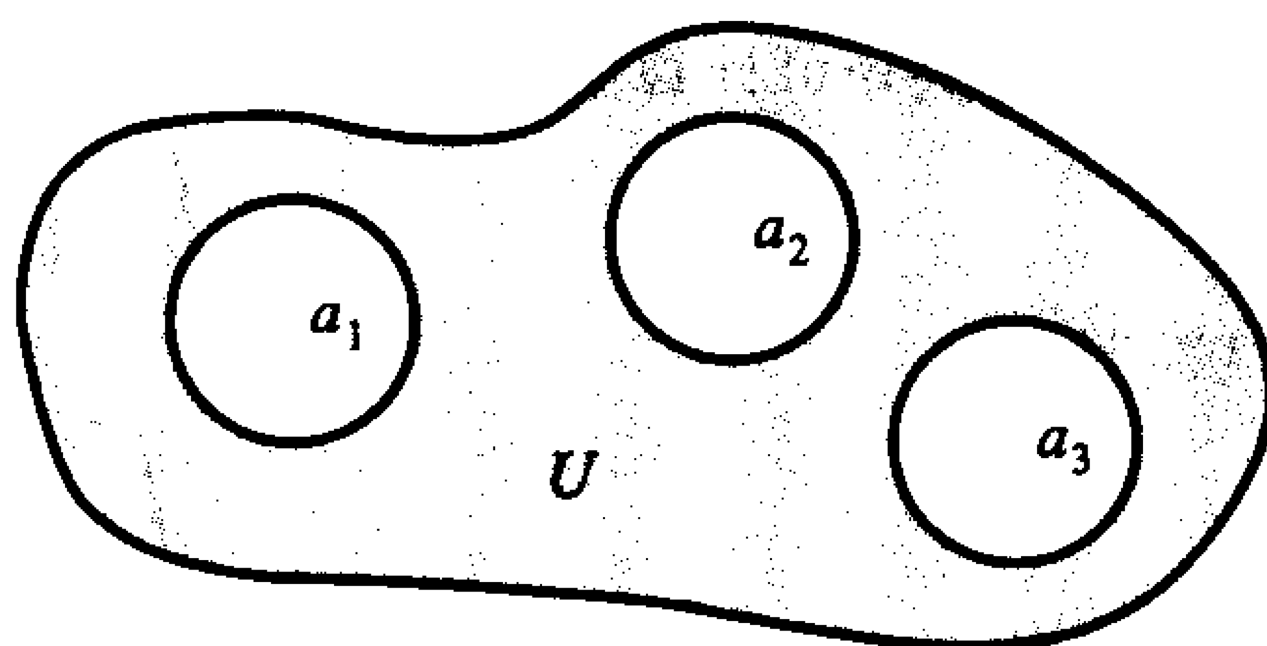


图 12.19

13.15 证明如果 f 在 U 中全纯, 则对任意 $\xi \in U$ 有

$$f(\xi) = \frac{1}{2\pi i}\int_{\partial U}\frac{f(z)}{z-\xi}dz.$$

13.16 设 f 在 $U - \{a_1\}\bigcup\cdots\bigcup\{a_k\}$ 中全纯 (并且 f 在 $W - \{a_1\}\bigcup\cdots\bigcup\{a_n\}$ 中为光滑, 其中 $W \supset \overline{U}$). 假设在每个 a_i 的某个邻域中

$$f(z) = B_{h_i,i}(z-a_i)^{-hi} + \cdots + B_{1,i}(z-a_i)^{-1} + \varphi_i(z),$$

其中 $\varphi_i(z)$ 在此邻域中为全纯 (见图 12.19). 证明

$$\frac{1}{2\pi i}\int_{\partial U} f(z)dz = B_{11} + B_{12} + \cdots + B_{1k}.$$

[提示: 把习题 13.13 用于 $U - \bigcup D_{i,\varepsilon}$, 其中 $D_{i,\varepsilon}$ 为围绕 a_i 的小圆盘, 然后让 $\varepsilon \to 0$.]

习题 13.16 的结果常常被用做计算定积分. 例如, 假设我们想要计算

$$\int_0^{2\pi} R(\cos\theta, \sin\theta)d\theta,$$

其中 R 为两个变量的某个有理函数. 如果我们令 $z = e^{i\theta}$, 则此积分成为

$$-i\int_{\partial D_1} R\left[\frac{1}{2}(z+z^{-1}), \frac{1}{2i}(z-z^{-1})\right]\frac{dz}{z},$$

其中 D_1 是单位圆盘. 要运用习题 13.15, 我们只须找出这些 a_j 点及系数 $B_{1,j}$ 来计算此积分. 例如, 如果 $a > 0$, 则

$$\int_0^{\pi}\frac{d\theta}{a+\cos\theta} = \frac{1}{2}\int_0^{2\pi}\frac{d\theta}{a+\cos\theta} = -i\sum -i\int_{\partial D_1}\frac{dz}{z^2+2az+1}.$$

现在 $z^2+2az+1 = (z-a_1)(z-a_2)$, 其中 $a_1 = -a+(a^2-1)^{1/2}, a_2 = -a-(a^2-1)^{1/2}$. 因此只有 $a_1 \in D_1$, 而且

$$\frac{1}{z^2+2az+1} = \frac{1}{a_1-a_2}\left(\frac{1}{z-a_1} - \frac{1}{z-a_2}\right),$$

故此积分可计算如下:

$$\int_0^{\pi} \frac{d\theta}{a+\cos\theta} = \frac{\pi}{\sqrt{a^2-1}}.$$

13.17 计算

$$\int_0^{\pi/2} \frac{d\theta}{a+\sin^2\theta}(a>0).$$

13.18 计算

$$\int_0^{2\pi} \frac{\sin^2\theta}{a+b\cos\theta}d\theta.$$

设 P 和 Q 为多项式使得 $Q(x)\neq 0$ 对所有实数 x 成立. 假设

$$P(z) = a_{n-2}z^{n-2} + a_{n-3}z^{n-3} + \cdots + a_1z + a_0$$

和

$$Q(z) = b_nz^n + b_{n-1}z^{n-1} + \cdots + b_1z + b_0,$$

其中 $b_n \neq 0$. 换句话说, Q 的次数至少要比 P 的次数大 2. 于是 P/Q 绝对可积, 且

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)}dx = \lim_{R\to\infty} \int_{-R}^{R} \frac{P(x)}{Q(x)}dx.$$

现在考虑 $\int_{\partial U}[P(z)/Q(z)]dz$, 其中 U 为半径为 R 的半圆盘 (见图 12.20). 在 ∂U 上的积分分裂为两部分:

$$\int_{-R}^{R} \frac{P(x)}{Q(x)}dx + \int_{\Gamma_R} \frac{P(z)}{Q(z)}dz,$$

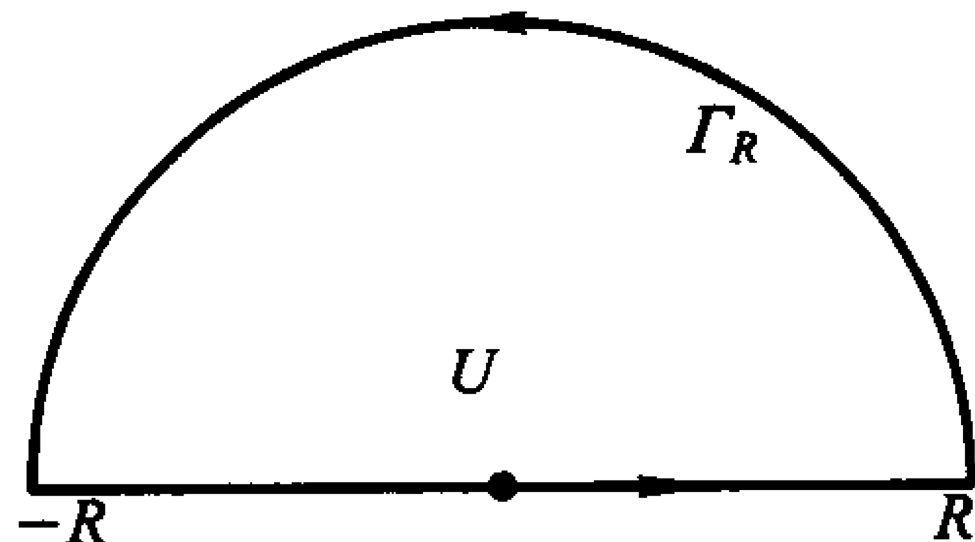


图 12.20

其中 Γ_R 是半径为 R 的半圆. 对大 R 值 $\|z\|=R$:

$$\left|\frac{P(z)}{Q(z)}\right| \leqslant \frac{1}{R^2}\frac{|a_{n-2}|+|a_{n-3}|\frac{1}{R}+\cdots+|a_0|\frac{1}{R^{n-2}}}{\left(|b_n|-|b_{n-1}|\frac{1}{R}-\cdots-|b_0|\frac{1}{R^n}\right)} < \frac{C}{R^2},$$

故

$$\left|\int_{\Gamma_R} \frac{P(z)}{Q(z)}dz\right| = \left|iR\int_0^{\pi} \frac{P(Re^{i\theta})}{Q(Re^{i\theta})}d\theta\right| \leqslant \frac{c\pi}{R},$$

那么有 $\lim_{R\to\infty}\int_{\Gamma_R}=0$. 于是我们可以应用习题 13.16 了, 因为这时 Q 在上半平面只有有限个零点.

13.19 计算

$$\int_{-\infty}^{\infty} \frac{dx}{(x^2+a^2)^2} \qquad (a>0).$$

13.20 证明

$$\int_{-\infty}^{\infty} \frac{dx}{(x^2+1)^n} = \frac{\pi}{2^{2(n-1)}}\frac{(2n-2)!}{[(n-1)!]^2}.$$

第十三章　经典力学

在这一章里我们要给出对经典理论力学的一个简短介绍. 要强调指出, 我们在学习的是一个数学分支: 在力学系统的物理所产生的许多问题中普遍存在的数学思考里得出的理想化描述. 因此我们要在公理基础上制定一个数学模型, 用来描述在研究力学方程中所产生的特征性质. 这个模型将用于力学应用中出现的大多数情形. 为了简化数学, 我们舍弃了某种广泛性, 因此在经典力学中的某些情形, 我们的模型并不合适.

力学致力于研究一个物理系统如何随时间演变. 第一个基本假定是说此系统可以 (局部地) 由一组连续参数描述. 更准确地说, 我们假定了此物理系统的所有可能的"位置"或构形是一个微分流形 M, 我们称其为*构形空间*.

举例说, 如果此物理系统由三个在空间自由运动的质点组成, 我们则可以用给出这三个质点在任何给定瞬间的位置来描述此系统的"位置". 因此, 这种情形的构形空间是

$$\mathbb{E}^3 \times \mathbb{E}^3 \times \mathbb{E}^3.$$

若坚持认为没有两个质点能在同一时刻占有同一个空间位置, 则构形空间 M 由

$$M = \mathbb{E}^3 \times \mathbb{E}^3 \times \mathbb{E}^3 - S$$

给出, 其中

$$S = \{\prec v_1, v_2, v_3 \succ : v_i \in \mathbb{E}^3, \text{且} v_1 = v_2 \text{或} v_1 = v_3 \text{或} v_2 = v_3\}.$$

如果此物理系统是一个绕某个空间固定点旋转的刚体 (譬如陀螺), 那么这个系统的位置就完全由在此物体中的三个法正交的向量 (譬如通过这个固定点所作的) 的空间位置所决定. 因为这三个向量可以放在任何位置而只要保持法正交并且不改变定向 (即从右手系变成左手系之类), 所以我们便看出 M 是 $\mathbb{E}^3$ 中所有定向法正交基全体的集合. 这时的 M 是个三维流形. 如果固定某个任意的初始法正交基 $\prec e_1, e_2, e_3 \succ$, 那么此系统的任意可能的位置由 $\prec v_1, v_2, v_3 \succ$ 给出, 其中 $v_i = Ae_i$, A 是个旋转, 即 A 是个具正行列式的正交线性变换. 因而 M 微分同胚于 $0^+(3)$, 即所有正行列式的 3×3 正交矩阵的空间.

力学的基本问题是描述系统的构形是如何随时间变化的. 当一个系统演变时, 在任意瞬时 t, 此构形由某个点 $C(t) \in M$ 给出. 我们的问题是要给出对此可能的曲线 C 一个合理而简明的描述, 而这条曲线 C 是在此力学系统的构形中的实际变化所产生的.

经典力学的第二个基本假定是曲线 $C(t)$ 可以被此系统在任何给定时间的 "状态" 的信息决定. 那就是说, 我们不得不 (一般来说, 应该) 了解在已知时刻关于此系统的比构形更多的东西, 以便能够预测它未来的构形. 然而, 如果我们确实有了足够的这类瞬时信息, 我们就可以决定出曲线 $C(t)$. 合适的信息总量被称作此系统的状态. 假设所有状态的集合本身是一个微分流形 $\mathcal{S}$. 因为一个系统的状态包含了比构形更多的信息, 我们便可以对每个状态 s 指定一个此系统在状态 s 的构形 $\pi(s) \in M$. 换句话说, 我们有映射 $\pi: \mathcal{S} \to M$. 假定映射 π 是可微的.

假设当我们知道了此系统在时刻 t_0 的状态 s 时我们便能预测出在任何未来时刻 t 的它的状态. 因而我们给出了一个映射

$$\varphi_{t,t_0}: \mathcal{S} \to \mathcal{S}$$

使得如果此系统在时刻 t_0 处于状态 s, 则它在时刻 t 的状态为 $\varphi_{t,t_0}(s)$. 现在关于时刻 t_0 便没有任何特殊之处了. 如果 $t > t_1 > t_0$, s 为在时刻 t_0 的状态, 则 $\varphi_{t_1,t_0}(s)$ 为在时刻 t_1 的状态, 因此

$$\varphi_{t,t_1}(\varphi_{t,t_0}(s))$$

便是在时刻 t 的状态.

换句话说,

$$\varphi_{t,t_1} \circ \varphi_{t_1,t_0} = \varphi_{t,t_0}$$

我们最终发现在大多数的 (基本的) 力学系统中, 如果适当地对时间参数化, 则函数 φ_{t,t_0} 真正地只依赖于 $t-t_0$.(这对于所谓的非守恒系统不再成立. 典型的非守恒系统是包含了摩擦力的力学系统. 这就是通常不去研究一个十分完整的系统的原因. 在有摩擦力的情形, 例如, 热量的损耗, 必须要把它考虑进去.) 让我们假定 φ_{t,t_0} 只依赖于 $t-t_0$. 那么, 如果我们记

$$s_1 = t_1 - t_0, s_2 = t - t_1,$$

则前面的方程式可以完成

$$\varphi_{s_2} \circ \varphi_{s_1} = \varphi_{s_1+s_2}.$$

除去我们至今仍限制于 s 的非负值这一点外, 这看起来像在定义一个单参数群. 事实上可以假定这个限制并非必要的, 我们实际上有了 $\mathcal{S}$ 上的一个流 φ.

再重复一遍. 我们有了表示我们系统的所有可能性的那个构形集合形成的一个微分流形 M, 一个代表此系统所有可能状态的微分流形 $\mathcal{S}$, 一个可微映射 $\pi:\mathcal{S}\to M$, 以及 $\mathcal{S}$ 上的一个流 φ. 于是曲线 $C(t)$ 全都被假定具有形式

$$C(t) = \pi\varphi_t(x),$$

其中的某个 $x\in\mathcal{S}$.

我们因此必须对任意给出的构形空间 M 描述 $\mathcal{S}$, π 和 φ, 经典力学关于空间 $\mathcal{S}$(及映射 π) 的性质作出了非常明确的假定. 它断定一个系统的状态完全被它的 (构形和它的) "动量" 决定. 什么是我们这种抽象架构的动量? 像通常那样, 动量应该是某种反抗速度变化的东西. 最后发现代表 "对瞬时速度变化的无穷小阻抗" 恰当的东西原来是一个余切向量. 对此的一个具有启发性的动机是这样的 (读者可选择略去不看.): 在任一个已知构形 $x\in M$ 上, 所有可能的速度向量的集合恰好是 $T_x(M)$. 在任意 $v\in T_x(M)$, 一个 "对 v 变化的阻抗" 可以是某个在 v 附近有定义而在 v 为零的函数. 在一阶范围内, 我们可以将这函数换成它的微分. 因此 "无穷小阻抗" 是 $T_v(T_x(M))$ 上的线性函数. 由于 $T_x(M)$ 是个向量空间, 我们对每个 v 可将 $T_v(T_x(M))$ 等同于 $T_x(M)$. 因而所有可能的 "动量" 等同于 $T_x^*(M)$ 中所有的元素.

集合 $\mathcal{S}$ 因此被当作所有动量的集合, 即在 M 上所有点的所有余切向量的集合. 我们必须首先指出如何把这个空间做成一个微分流形.

13.1 切丛和余切丛

设M 为微分流形, 我们来考虑在 M 的所有点上的所有切向量的集合 $T(M)$,

因而

$$T(M) = \bigcup_{x \in M} T_x(M).$$

设 π 表示 $T(M)$ 到 M 上的映射, 它指派给每个切向量为它所定义在那里的 M 的点; 就是说, 如果 $\xi \in \dot{T}_x(M)$, 则 $\pi(\xi) = x$. 我们断言 $T(M)$ 可以以一种自然的方式做成一个微分流形, 从而使 π 为可微映射. 事实上, 设 $\mathfrak{a}$ 为 M 上的一个总图表, 对每个 $(U, \alpha) \in \mathfrak{a}$, 定义 $T(\alpha) : \pi^{-1}(U) \to V \oplus V$ 为

$$T(\alpha)\xi = \prec \alpha(x), \xi_\alpha \succ, \tag{1.1}$$

其中 $\xi \in T_x(M), x \in U$. 我们断定集合 $\{(\pi^{-1}U, T(\alpha))\}$ 是 $T(M)$ 上的一个总图表. 要验证 $T(\alpha)$ 是一一的, 我们看出, 如果 $\xi \in T_x(M), \eta \in T_y(M)$, 其中 $x \neq y$, 则 $\alpha(x) \neq \alpha(y)$; 而如果 $\xi \neq \eta$ 同属于 $T_x(M)$, 我们则有 $\xi_\alpha \neq \eta_\alpha$. 要验证满足转移法则, 我们则注意到第九章的等式 (4.3) 表明如果 $(U, \alpha), (W, \beta)$ 为 $\mathfrak{a}$ 的两个分图表, 则

$$T(\beta) \circ T(\alpha)^{-1}(v, \xi_\alpha) = \prec \beta \circ \alpha^{-1}(U), J_{\beta \circ \alpha^{-1}}(v)\xi_\alpha \succ, \tag{1.2}$$

其中 $\prec v, \xi_\alpha \succ \in T(\alpha)\pi^{-1}(U \bigcap W)$, 即 $v \in \alpha(U \bigcap W)$, 而 ξ_α 在 V 中为任意.

$T(M)$ 上的这个结构不依赖于 $\mathfrak{a}$ 的选取这个事实是显然的. π 为可微也是清楚的; 事实上, $(\pi^{-1}(U), T(\alpha))$ 和 (U, α) 是相容的分图表, 且在它们的表示下, $\alpha \circ \pi \circ T(\alpha)^{-1}(v, \xi_\alpha) = v$. 我们称 $T(M)$ 连同它作为微分流形的结构为 M 的切丛.

如果 M 是有限维的, 设 (U, α) 为 M 的一个分图表, 其坐标为 $x^1, \cdots, x^n$, 从而

$$\alpha(x) = \prec x^1(x), \cdots, x^n(x) \succ \in \mathbb{R}^n.$$

我们表示在 $\pi^{-1}(U)$ 上与分图表 $T(\alpha)$ 相关的坐标为 $\prec q^1, \cdots, q^n, \dot{q}^1, \cdots, \dot{q}^n \succ$, 因此, 如果 $\xi \in T_x(M), x \in U$, 我们则有

$$\begin{aligned} T(\alpha)(\xi) &= \prec \alpha(x), \xi_\alpha \succ \\ &= \prec q^1(\xi), \cdots, q^n(\xi), \dot{q}^1(\xi), \cdots, \dot{q}^n(\xi) \succ, \end{aligned}$$

从而

$$q^i = x^i \circ \pi, \quad \xi = \sum \dot{q}^i(\xi) \left(\frac{\partial}{\partial x^i} \right)_x. \tag{1.3}$$

换句话说, q^i 正好是坐标 x^i, 但通过 π 被看做 $T(M)$ 上的函数, $\dot{q}^i(\xi)$ 是 ξ 关于基

$$\left\{ \left(\frac{\partial}{\partial x^1} \right)_x, \cdots, \left(\frac{\partial}{\partial x^n} \right)_x \right\}$$

的分量.

我们可以按照相似的步骤对

$$T^*(M)=\bigcup_{x\in M}T_x^*(M)$$

进行论证, 从而得到了余切丛. 如果 (U,α) 是 M 上的总图表 $\mathfrak{a}$ 的一个分图表, 我们定义映射

$$T^*(\alpha):\pi^{-1}(U)\to V\oplus V^*$$

为

$$T^*(\alpha)(l)=\prec\alpha(x),l_\alpha\succ, \tag{1.4}$$

其中 $l\in T_x^*(M),x\in U$. 如同前面那样, 它定义了 $T^*(M)$ 上的一个总图表, 又定义了 $T^*(M)$ 上的一个可微结构, 其与 M 上的总图表选取无关.(注意, 我们已经用了同一个字母 π 来表示两个射影. $T(M)\to M$ 和 $T^*(M)\to M$. 一旦有混淆之虞, 我们将把它们分别以 π_T 和 π_{T^*} 表示.)

如果 M 是有限维, (U,α) 为具坐标 $x^1,\cdots,x^n$ 的分图表, 我们将用

$$\prec q^1,\cdots,q^n,p^1,\cdots,p^n\succ$$

表示 $(\pi^{-1}(U),T^*(\alpha))$ 的坐标. 因此, 如果 $l\in T_x^*(M),x\in U$, 我们则有

$$T^*(\alpha)l=\prec\alpha(x),l_\alpha\succ=\prec q^1(l),\cdots,q^n(l),p^1(l),\cdots p^n(l)\succ,$$

故

$$q^i=q^i\circ\pi,\quad l=\sum p^i(l)(dx^i)_x. \tag{1.5}$$

13.2 变分方程

设 M_1 和 M_2 为两个微分流形, φ: $M_1\to M_2$ 为可微映射. 于是 φ 诱导了一个映射 $T(\varphi):T(M_1)\to T(M_2)$, 定义为, 对 $\xi\in T_x(M_1)$,

$$T(\varphi)\xi=\varphi_{*x}(\xi). \tag{2.1}$$

为验证 $T(\varphi)$ 可微, 我们选择相容的 M_1 上的分图表 (U,α) 和 M_2 上的 (W,β) . 那么由 $T(\varphi)$ 的定义知道

$$(T(U),T(\alpha))\quad 和\quad (T(W),T(\beta))$$

为相容的分图表. 另外

$$T(\beta)\circ T(\varphi)\circ T(\alpha)^{-1}(v_1,v_2)=\prec(\beta\circ\varphi\circ\alpha^{-1})v_1,J_{\beta\circ\varphi\circ\alpha^{-1}}(v_1)v_2\succ. \tag{2.2}$$

这样便建立了 $T(\varphi)$ 的可微性. 我们也看到, 如果 ψ: $M_2\to M_3$ 是第二个可微映射, 则由第九章等式 (4.4) 知

$$T(\psi\circ\varphi)=T(\psi)\circ T(\varphi). \tag{2.3}$$

另外, 由定义清楚地知道

$$\pi\circ T(\varphi)=\varphi\circ\pi. \tag{2.4}$$

换句话说, 图表

$$\begin{array}{ccc} T(M_1) & \xrightarrow{T(\varphi)} & T(M_2) \\ \pi\downarrow & & \downarrow\pi \\ M_1 & \underset{\varphi}{\longrightarrow} & M_2 \end{array}$$

交换, 它的意思是选取从 $T(M_1)$ 到 M_2 的随便哪一条路线都没关系.

特别, 如果 $\{\varphi_t\}$ 是 M 的一个单参数微分同胚群, 我们则得到 $T(M_1)$ 上的一个单参数群 $T(\varphi_1)$, 由 (2.4) 式, 其中

$$\pi\circ T(\varphi_t)\xi=\varphi_t(\pi(\xi)). \tag{2.5}$$

如果 X 是 $\{\varphi_t\}$ 的无穷小生成元, 我们则以 $T(X)$ 表示 $T(\varphi_t)$ 的无穷小生成元. 由 (2.5) 式知, 如果 $\xi\in T_x(M)$, 则

$$\pi_{*\xi}(T(X)_\xi)=X_x. \tag{2.6}$$

让我们在一个分图表中得出 $T(X)$ 的表达式: 设 (U,α) 为一分图表, 选择 W 为满足 $\overline{W}\subset U$ 的开集, 而 $\varepsilon>0$ 使得对 $|t|<\varepsilon$ 有 $\varphi_t(W)\subset U$. 于是

$$T(\alpha)\circ T(\varphi_t)\circ(T(\alpha))^{-1}\prec v,w\succ=\prec(\alpha\circ\varphi_t\circ\alpha^{-1})v,J_{(\alpha\circ\varphi_t\circ\alpha^{-1})}(v)w\succ,$$

故, 在 $t=0$ 取关于 t 的微分, 由第九章的等式 (4.9) 有

$$T(X)_{T(\alpha)}\prec v,w\succ=\prec X_\alpha(v),dX_{\alpha(v)}(w)\succ. \tag{2.7}$$

如果 M 是有限维的, $x^1,\cdots,x^n$ 为 (U,α) 的坐标, 而

$$X_\alpha=\prec X^1,\cdots,X^n\succ,$$

我们则可重写 (2.7) 式为

$$T(X)_{T(\alpha)} \prec v, w \succ = \prec X^1(v), \cdots, X^n(v), \sum \frac{\partial X^1}{\partial x^j} w^j, \cdots, \sum \frac{\partial X^n}{\partial x^j} w^j \succ, \quad (2.7')$$

其中 $w = \prec w^1, \cdots, w^n$. 换句话说, 利用局部坐标 $\prec q^1, \cdots, q^n, \dot{q}^1, \cdots, \dot{q}^n \succ$, 对应于 $T(X)$ 的微分方程有形式

$$\frac{dq^i}{dt} = X^i(q^1, \cdots, q^n) \tag{2.8}$$

和

$$\frac{d\dot{q}^i}{dt} = \frac{\partial X^i}{\partial x^1}\dot{q}^1 + \cdots + \frac{\partial X^i}{\partial x^n}\dot{q}^n. \tag{2.9}$$

注意方程 (2.8), 它正好是对应于向量场 X 的局部方程. 因为 $q^i = x^i \circ \pi$, 这只不过是 (2.6) 式的另一种写法. 假设

$$\varphi_t(x) = \prec x^1(t), \cdots, x^n(t) \succ$$

是 X 在 U 中的解曲线. 于是我们可以把系数

$$\frac{\partial X^i}{\partial x^j} = \frac{\partial X^i}{\partial x^j}(x^1(t), \cdots, x^n(t)),$$

看为只是 t 的函数. 因而 (2.9) 式有形式

$$\frac{dw}{dt} = A(t)w,$$

它是 w 的线性微分方程. 称此线性微分方程为 X 沿曲线 $\varphi(x)$ 的**变分方程**. 粗略地说, 从线性逼近来看, 它代表了在 $\varphi(x)$ 附近的 X 的解曲线是如何偏离 $\varphi(x)$ 的.

现在我们来完成对余切丛的类似构造. 如果 φ: $M_1 \to M_2$ 为可微映射, 则 $\varphi_x^* : T^*_{\varphi(x)}(M_2) \to T^*_x(M_1)$ 的进行方向不对头. 因此我们把注意力局限于局部微分同胚的那些映射, 从而定义 $T^*(\varphi)$ 为, 对 $l \in T^*_x(M_1)$,

$$T^*(\varphi)l = (\varphi^{-1})^* l = (\varphi^*_{\varphi(x)})^{-1} l. \tag{2.10}$$

因为 φ 为微分同胚, 所以

$$(\varphi^*)^{-1} : T^*_x(M_1) \to T^*_{\varphi(x)}(M_2)$$

有确切定义, 并且有

$$\pi\circ T^*(\varphi)=\varphi\circ\pi. \tag{2.11}$$

如果 (U,α) 和 (W,β) 为 M_1 和 M_2 上的相容分图表, 则

$$T^*(\beta)\circ T^*(\varphi)\circ T^*(\alpha)^{-1}\prec v_1,v_2\succ=\prec(\beta\circ\varphi\circ\alpha^{-1})v_1,[J_{\beta\circ\varphi\circ\alpha^{-1}}(v_1)]^{*-1}v_2\succ, \tag{2.12}$$

从而 $T^*(\varphi)$ 可微.

如果 φ: $M_1\to M_2,\psi$: $M_2\to M_3$ 为微分同胚, 则

$$T^*(\psi\circ\varphi)l=(\psi\circ\varphi)^{*-1}l=(\varphi^*\circ\psi^*)^{-1}l=\psi^{*-1}\circ\varphi^{*-1}l,$$

从而

$$T^*(\psi\circ\varphi)=T^*(\psi)\circ T^*(\varphi). \tag{2.13}$$

特别地, 如果 $\{\varphi_t\}$ 是流形 M 上的一个流, 我们则得到 $T^*(M)$ 上的一个流 $T^*(\varphi_t)$. 它满足

$$\pi\circ T^*(\varphi_t)l=\varphi_t(\pi(l)). \tag{2.14}$$

如果 X 为 $\{\varphi_t\}$ 的无穷小生成元, 则以 $T^*(X)$ 表示 $\{T^*(\varphi_t)\}$ 的无穷小生成元. 它满足

$$\pi_*T^*(X)_l=X_x, \tag{2.15}$$

其中 $l\in T^*_x(M)$.

13.3 $T^*(M)$ 上的基本线性微分形式

在回到力学的研讨之前, 我们要较详细地研究一下余切丛的几何. 设 M 为微分流形, z 为 $T^*(M)$ 的点. 设 ξ 为在点 z 切于 $T^*(M)$ 的切向量, 故

$$\xi\in T_z(T^*(M)).$$

于是 $\pi_*\xi$ 是 $T_{\pi(z)}(M)$ 中的一个元素. 因为 $z\in T^*_{\pi(z)}(M)$ 是 $T_{\pi(z)}(M)$ 上的线性函数, 我们可以考虑表达式

$$\langle\pi_*\xi,z\rangle,$$

它线性地依赖于 ξ. 把这个 ξ 的线性函数记为 θ_z. 我们因此定义了 $T^*(M)$ 上一个线性微分形式:

$$\langle\xi,\theta_z\rangle=\langle\pi_{*z}\xi,z\rangle,\qquad \text{其中}\qquad \xi\in T_z(T^*(M)). \tag{3.1}$$

称形式 θ 为 $T^*(M)$ 的**基本线性形式**. 我们来得出 θ 在一个分图表 $(\pi^{-1}(U), T^*(\alpha))$ 下的表达式. 因为 $T^*(\alpha)$ 把 $\pi^{-1}(U)$ 映到 $V\oplus V^*$ 的一个开集 O 上, 所以 $\theta_{T^*(\alpha)}$ 的表达式应该是从 O 到 $(V\oplus V^*)^*$ 的一个函数, 而后者可以等同于 $V^*\oplus V$. 让我们计算这个函数. 在 M 的分图表 (U,α) 和 $T^*(M)$ 上的 $(\pi^{-1}(U),T^*(\alpha))$ 下, 我们有

$$\xi_{T^*(\alpha)} =\prec v, w^* \succ \in V \oplus V^*, (\pi_*\xi)_\alpha = v \in V,$$

故

$$\langle \pi_*\xi, z\rangle = \langle v, z_\alpha\rangle.$$

因此 $\langle \prec v, w^* \succ, (\theta_z)_{T^*(\alpha)}\rangle = \langle v, z_\alpha\rangle$; 就是说,

$$(\theta_z)_{T^*(\alpha)} =\prec z_\alpha, 0 \succ \in V^*\oplus V = (V \oplus V^*)^*.$$

换句话说, 在分图表 $T^*(\alpha)$ 下微分形式 θ 的局部表达式是由

$$\theta_\alpha \prec u_1, u_2^* \succ =\prec u_2^*, 0 \succ \tag{3.2}$$

给出. 如果 M 为有限维, 并且使用局部坐标 $\prec q^1,\cdots,q^n,p^1,\cdots,p^n \succ$, 则

$$\pi_*\left(\frac{\partial}{\partial q^i}\right)_z = \frac{\partial}{\partial x^i},$$

而

$$\pi_*\left(\frac{\partial}{\partial p^i}\right)_z = 0, z = \sum p^i(z)dx_x^j.$$

因此

$$\left\langle\left(\frac{\partial}{\partial q^i}\right)_z, \theta_z\right\rangle = \left\langle \frac{\partial}{\partial x^i}, \sum p^j(z)dx_x^j\right\rangle = p^i(z),$$

而

$$\left\langle\left(\frac{\partial}{\partial p^i}\right)_z, \theta_z\right\rangle = 0,$$

从而

$$\theta_z = \sum p^i(z)dq_z^i$$

或者

$$\theta = \sum p^i dq^i. \tag{3.3}$$

自然, (3.3) 式正好是由 $V=\mathbb{R}^n$ 的一个基下写出 (3.2) 式的一种方式.

设 $\varphi: M_1 \to M_2$ 为微分同胚, θ_1 为 $T^*(M_1)$ 上的基本线性形式, 又设 θ_2 是 $T^*(M_2)$ 上的基本线性形式. 设

$$\xi \in T_z(T^*(M_1)), \pi(z) = x \in M_1.$$

于是有

$$\pi_* T^*(\varphi)_* \xi = \varphi_* \pi_* \xi,$$

从而

$$\begin{aligned}\langle T^*(\varphi)_*\xi, \theta_{2T^*(\varphi)z}\rangle &= \langle \varphi_{*x}\pi_*\xi, T^*(\varphi)z\rangle \\ &= \langle \varphi_{*x}\pi_*\xi, (\varphi^*)^{-1}_{\varphi(x)}z\rangle \\ &= \langle \pi_*\xi, z\rangle = \langle \xi, \theta_{1z}\rangle.\end{aligned}$$

换句话说,

$$(T^*(\varphi))^*\theta_2 = \theta_1. \tag{3.4}$$

特别地, 如果 $\{\varphi_t\}$ 是 M 上的一个流, 则

$$(T^*(\varphi_t))^*\theta = \theta.$$

如果 X 是 M 上的一个向量场, 那么上面最后面的那个方程的无穷小形式便是

$$D_{T^*(x)}\theta = 0. \tag{3.5}$$

注意, M 上的任意一个向量场 X 定义了 $T^*(M)$ 上一个函数 f_X 为: 当 $\pi(z) = x$ 时

$$f_X(z) = \langle X_x, z\rangle. \tag{3.6}$$

我们也有

$$f_X = \langle T^*(X), \theta\rangle, \tag{3.7}$$

这是因为由 (2.15) 式 $\langle T^*(X)_z, \theta_z\rangle = \langle \pi_* T^*(X)_z, z\rangle = \langle X_x, z\rangle$.

最后, 由 (3.5) 式和第十一章的等式 (6.14), 我们有

$$0 = D_{T^*(X)}\theta = d\langle T^*(X), \theta\rangle + T^*(X)\rfloor d\theta,$$

故

$$df_X = -T^*(X)\rfloor d\theta. \tag{3.8}$$

由于稍后才能明白的理由, 有时称函数 f_X 为与向量场 X 相关联的动量函数.

13.4 $T^*(M)$ 上的基本外 2 - 形式

最终可证明, 外 2 - 形式

$$\Omega = d\theta \tag{4.1}$$

在力学中起了基本性的作用, 因此我们要研究它的一些性质. 首先, 由于 $d^2 = 0$, 我们有

$$d\Omega = 0. \tag{4.2}$$

我们断言 Ω_z 是 $T_z(T^*(M))$ 上一个非异双线性形式, 其中 $z \in T^*(M)$ 为任意. 就是说,

$$\text{如果}\ \ \xi{\in}T_z(T^*(M))\ \ \text{使得}\ \ \xi\rfloor\Omega_z = 0,\ \ \text{则}\ \xi = 0. \tag{4.3}$$

为此, 我们按照分图表 $(\pi^{-1}(U), T^*(\alpha))$ 计算 Ω 的局部表达式 [其中 $z \in \pi^{-1}(U)$]. 映射 $T^*(\alpha)$ 给出了 $\pi^{-1}(U)$ 与 $V \oplus V^*$ 的一个子集的微分同胚, 从而由 (3.4) 式, M 上的形式 θ 转成了在 $V \oplus V^* = T^*(V)$ 上的对应形式 θ_α. θ_α 也可以被看为一个由 (3.2) 给出的 $(V^* \oplus V)$ - 值函数. 我们以 ξ_α 表示 $\xi_{T^*(\alpha)}$, 并把

$$\xi_\alpha = \prec X_1, X_2 \succ \in V \oplus V^*$$

想成是 $V \oplus V^*$ 上的一个常值向量场. 于是

$$\langle \xi_\alpha, \theta_\alpha \rangle_{\prec u_1, u_2^* \succ} = \langle X_1, u_2^* \rangle,$$

从而 $d\langle \xi_\alpha, \theta_\alpha \rangle$ 是个线性微分形式, 它由

$$\langle \eta_\alpha, d\langle \xi_\alpha, \theta_\alpha \rangle \rangle = \langle X_1, Y_2 \rangle,\ \text{其中}\ \eta_\alpha = \prec Y_1, Y_2 \succ \in V \oplus V^*.$$

另一方面, 由于 ξ_α 是个常值向量场, 故而李导数 $D_{\xi_\alpha}\theta_\alpha$ 化成线性函数 $\prec u_2, u_2^* \succ \mapsto \prec u_2^*, 0 \succ$ 的通常导数. 此导数正是常值 $\prec X_2, 0 \succ$. 因此李导数 $D_{\xi_\alpha}\theta_\alpha$ 由常值线性形式

$$D_{\xi_\alpha}\theta_\alpha = \prec X_2, 0 \succ \in V^* \oplus V$$

给出, 就是说,

$$\langle \eta_\alpha, D_{\xi_\alpha}\theta_\alpha \rangle = \langle Y_1, X_2 \rangle.$$

现在 $\xi_\alpha \rfloor \Omega_\alpha = \xi_\alpha \rfloor d\theta_\alpha = D_{\xi_\alpha}\theta_\alpha - d\langle \xi_\alpha, \theta_\alpha \rangle$; 所以我们看出 $\xi_\alpha \rfloor \Omega_\alpha$ 为常值形式

$$\xi_\alpha \rfloor d\theta = \prec X_2, -X_1 \succ \in V^* \oplus V.$$

扼要地重复一下: 如果 $\xi \in T_z(T^*(M))$ 使得 $\xi_{T^*(\alpha)}$ 由 $\xi_{T^*(\alpha)} = \prec X_1, X_2 \succ \in V \oplus V^*$ 给出, 则

$$(\xi \rfloor \Omega_z)_{T^*(\alpha)} = \prec X_2, -X_1 \succ \in V^* \oplus V. \tag{4.4}$$

等式 (4.3) 显然由此推出.

由于 (4.3) 式具有基本的重要性, 我们给出在有限维流形情形下的另一个不同的推导. 如果我们用坐标 $\prec q^1, \cdots, q^n, p^1, \cdots, p^n \succ$, 则

$$\theta = \sum p^i dq^i,$$

从而

$$\Omega = \sum dp^i \wedge dq^i. \tag{4.5}$$

如果

$$X = \sum \left(A_i \frac{\partial}{\partial q^i} + B^i \frac{\partial}{\partial p^i} \right),$$

则

$$X \rfloor \Omega = \sum (B^i dq^i - A^i dp^i). \tag{4.6}$$

因此 $(X \rfloor \Omega)_z = 0$ 当且仅当 $X_z = 0$. 这证明了在 $T^*(M)$ 上我们有一个一一对应, $X \mapsto X \rfloor \Omega$, 将向量场对应到线性微分形式. 我们用 ω_X 表示 X 所对应的形式, 从而

$$\omega_X = X \rfloor \Omega;$$

并且以 X_ω 表示线性微分形式 ω 对应的向量场. 因而

$$\omega = X_\omega \rfloor \Omega = \omega_{X_\omega}.$$

我们看到有

$$d\omega = 0 \text{ 当且仅当 } D_{X_\omega} \Omega = 0. \tag{4.7}$$

事实上由 (4.2) 式,

$$D_{X_\omega} \Omega = d(X_\omega \rfloor \Omega) = d\omega.$$

特别地, 在 $T^*(M)$ 上有一个特定的向量场的类, 即对应于函数的, 也就是形如 X_{dF} 的向量场, 其中 F 是 $T^*(M)$ 上的函数. 称这个向量场为哈密顿 向量场.

* 由于有 (4.6) 式, 如果 $D_X \Omega = 0$, 则 X 至少局部地形如 X_{dF}, 这是因为任意满足 $d\omega = 0$ 的形式 ω 都可局部地写为 $\omega = dF$. 如果我们做一个拓扑假定,

说 $T^*(M)$ 的第一上同调群为 0, 那么 $D_X\Omega = 0$ 便等价于 X 为哈密顿. 这个假设是对构形空间 M 的性质的一个实质性的限制. 由于我们不想按这种方式去限制 M, 故我们不以 $D_X\Omega = 0$ 作为哈密顿向量场的定义. *

注意, 如果 X, Y 为哈密顿向量场, 则 $aX + bY$ 也是, 其中 a, b 为常数. 事实上, 如果 $X = X_{dF}, Y = X_{dG}$, 则

$$aX + bY = X_{d(aF+bG)}.$$

另外, $[X, Y]$ 也是哈密顿向量场. 事实上, 因为 $D_X\Omega = 0$, 有

$$D_X(Y\rfloor\Omega) = D_XY\rfloor\Omega + Y\rfloor D_X\Omega = D_XY\rfloor\Omega.$$

由于 $D_XY = [X, Y]$, 我们看出

$$[X, Y]\rfloor\Omega = D_X dG = dD_X G.$$

换句话说, 我们有

$$[X_{dF}, X_{dG}] = X_{d(X_{dF}G)}. \tag{4.8}$$

因此我们看到, 对应于哈密顿向量场的李括号我们有一个二元运算. 称其为**泊松括号**, 并以 $\{F, G\}$ 表示. 换而言之, 我们定义

$$\{F, G\} = X_{dF}G, \tag{4.9}$$

于是可重写 (4.8) 式为

$$[X_{dF}, X_{dG}] = X_{d\{F,G\}}. \tag{4.8'}$$

注意到

$$X_{dF}G = \langle X_{dF}, dG\rangle = \langle X_{dF}, X_{dG}\rfloor\Omega\rangle = \langle X_{dG}\wedge X_{dF}, \Omega\rangle \tag{4.10}$$

故而, 特别地,

$$\{F, G\} = -\{G, F\}. \tag{4.11}$$

利用局部坐标 $\prec q^1, \cdots, q^n, p^1, \cdots, p^n \succ$, 我们有

$$dF = \sum\left(\frac{\partial F}{\partial q^i}dq^i + \frac{\partial F}{\partial p^i}dp^i\right)$$

故由 (4.4) 式有

$$X_{dF} = \sum\left(\frac{\partial F}{\partial q^i}\frac{\partial}{\partial p^i} - \frac{\partial F}{\partial p^i}\frac{\partial}{\partial q^i}\right), \tag{4.12}$$

因此

$$\{F,G\}=\sum\left(\frac{\partial F}{\partial q^i}\frac{\partial G}{\partial p^i}-\frac{\partial F}{\partial p^i}\frac{\partial G}{\partial q^i}\right), \tag{4.13}$$

由此看出 (4.11) 式的反称性是显见的.(4.11) 式的一个推论是:

命题 4.1 *如果 F 和 G 是 $T^*(M)$ 上的函数, 使得 $X_{dF}G=0$ 则 $X_{dG}F=0$.*

换句话说, 如果 G 沿 X_{dF} 的解曲线为常值, 则 F 沿 X_{dG} 的解曲线也为常值.

事实上,

$$X_{dF}G=\{F,G\}=-\{G,F\}=-X_{dG}F,$$

命题 4.1 原来是力学的所有"守恒定律"的原型.

我们以下面的观察作为本节的结束. 设 Y 为 M 上的向量场. 那么 Y 的动量函数是 $T^*(M)$ 上一个函数 f_Y. 等式 (3.8) 断言

$$-T^*(Y)=X_{df_Y}. \tag{4.14}$$

13.5 哈密顿力学

如同我们在引言中表示的, 力学的第一基本假定是系统的演变由 $T^*(M)$ 上的一个流决定, 其中的 M 是此系统的构形空间. 第二基本假定是关于这个流的特性的. 它假定这个流的无穷小生成元是个哈密顿向量场. 那就是说, 假定存在 $T^*(M)$ 上的一个函数 H(称作能量)使得向量场 X_{-dH} 为 $T_*(M)$ 上描述这个系统演变的流的无穷小生成元.(负号是由于某个标准约定的结果.) 为了明白这个意思, 让我们把这个运动方程用有限维系统的 q 和 p 坐标表达出来. 因而有

$$X_{-dH}=\sum\left(\frac{\partial H}{\partial p^i}\frac{\partial}{\partial q^i}-\frac{\partial H}{\partial q^i}\frac{\partial}{\partial p^i}\right).$$

从而, 如果 $\prec q^1(\cdot),\cdots,q^n(\cdot),p^1(\cdot),\cdots,p^n(\cdot)\succ$ 是这个流的积分曲线, 那么它必定满足微分方程

$$\frac{dq^i}{dt}=\frac{\partial H}{\partial p^i},\qquad \frac{dp^i}{dt}=-\frac{\partial H}{\partial q^i}. \tag{5.1}$$

(4.11) 式的一个平凡推论是 $X_{-dH}H=0$. 换句话说, 函数 H 沿此系统的轨线为常数. 这个原理被称作*能量守恒定律*.

更一般地, 我们可将命题 4.1 阐述为

命题 5.1 设 X_{-dH} 是一个具能量 H 的力学系统的无穷小生成元. 设 F 是个函数, 满足

$$X_{dF}H = 0.$$

于是 F 在轨线上为常数, 这里说的轨线即由 X_{-dH} 生成的流的解曲线.

命题 5.1 是整个 "动量守恒" 定律的原型, 我们要在后面推导这个定律. 例如可参看在第 6 节末尾的讨论.

为了完全确定力学系统, 必须要给出函数 H. 能量原来是 (在许多但非全部情形) 两项之和, $H = K + U$, 其中的 K 被称作动能, 而称 U 为势能. 它们中每个都具有特殊的形式, 现在我们来描述它.

动能是 $T^*(M)$ 上的函数, 它与 M 上一个黎曼度量相关联. 设 $(\ ,\ )$ 是 M 上的一个黎曼度量. 它在每个 $T_x(M)$ 上给出了内积 $(\ ,\ )_x$, 因此诱导了 $T_x(M)$ 与 $T_x^*(M)$ 之间的一个同构, 从而给出了 $T_x^*(M)$ 上的内积, 我们仍记它为 $(\ ,\ )$. 函数 K 于是定义为

$$K(l) = \frac{1}{2}(l, l) \tag{5.2}$$

要了解黎曼度量与力学之间的重要关联, 让我们考虑最初等的情形: 研究 $\mathbb{E}^3$ 中质量 m 的单个质点. 速度和动量

$$\boldsymbol{p} = m\dot{\boldsymbol{q}}$$

之间的通常的关系可以详细阐述于后: 考虑 $\mathbb{E}^3$ 上的黎曼度量, 它为 $m\times$ 欧几里得度量. 就是说, 如果 $\prec x, y, z \succ$ 为 $\mathbb{E}^3$ 上的直角坐标, $\prec q_x, q_y, q_z, \dot{q}_x, \dot{q}_y, \dot{q}_z \succ$ 为 $T(\mathbb{E}^3)$ 上的对应坐标, 则

$$\|(\dot{q}_x, \dot{q}_y, \dot{q}_z)\|^2 = m\dot{q}_x^2 + m\dot{q}_y^2 + m\dot{q}_z^2.$$

于是 $T_x(\mathbb{E}^3) \to T_x^*(\mathbb{E}^3)$ 的映射由下面关系给出:

$$\prec q_x, q_y, q_z, \dot{q}_x, \dot{q}_y, \dot{q}_z \succ \mapsto \prec q_x, q_y, q_z, p_x, p_y, p_z \succ,$$

其中

$$p_x = m\dot{q}_x, p_y = m\dot{q}_y, p_z = m\dot{q}_z. \tag{5.3}$$

并且

$$K(q_x, q_y, q_z, p_x, p_y, p_z) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2). \tag{5.4}$$

因此从速度到动量的迁移依赖于一个黎曼度量的选取, 这个黎曼度量决定了从 $T(M)$ 到 $T^*(M)$ 的一个映射. 这可以被看做对构形的广义的“质量选择”.

函数 U 被假定具有形式 $U=\overline{U}\circ\pi$, 其中 $\overline{U}$ 是 M 上的函数. 形式 $\boldsymbol{F}=-d\bar{U}$ 被称作力场, 它的位势是 $\overline{U}$. 由于 M 上的黎曼度量, 它可以被视为 M 上的向量场: 对每个 $\xi\in T_x(M)$

$$(\xi,\boldsymbol{F}_x)=-\langle\xi,d\overline{U}\rangle.$$

在 $\mathbb{E}^3$ 中具质量 m 的单个质点的特殊情形中, 将 (5.3) 和 (5.4) 式代入方程 (5.1) 式, 在 $H=K+U$ 时, 给出了

$$\frac{dmq^i}{dt}=p^i,\qquad \frac{dp^i}{dt}=F^i,\tag{5.5}$$

或者, 因为 m 为常数,

$$m\frac{d^2q^i}{dt^2}=F^i,\tag{5.6}$$

这便是平常所说的力 = 质量 × 加速度的规则.

回到一般理论, 现在阐述命题 5.1 的一个有用的推论.

命题 5.2 设 $H=K+U$ 为与黎曼度量 (,) 关联的能量, 以及 M 上的函数 $\overline{U}$. 设 Y 为 M 上的向量场, 它是 (,) 的一个无穷小等距映射, 并且使得 $Y\overline{U}=0$. 于是动量函数 f_Y 在由 X_{-dH} 生成的流下为常数.

设 $\{\varphi_t\}$ 为 Y 生成的流. 因为 φ_t 为局部等距, $(\varphi_t^*l,\varphi_t^*l)=(l,l)$, 只要 φ_t^*l 有定义就成立, 从而

$$K(T^*(\varphi_t^*)l)=K(l)$$

因此

$$T^*(Y)K=0.$$

还有

$$\begin{aligned}T^*(Y)U&=\langle T*(Y),dU\rangle=\langle T*(Y),\pi^*d\overline{U}\rangle\\&=\langle\pi_*T^*(Y),d\overline{U}\rangle=\langle Y,d\overline{U}\rangle=0,\end{aligned}$$

使得 $T^*(Y)H=0$, 这样便可应用命题 5.1 [见等式 (4.14)].

13.6 中心力问题

在进行一般理论的讨论之前, 我们在一个简单而重要的情形中来解释前面所得到的结果. 我们要研究 $\mathbb{E}^3$ 中一个质点在一个 "集中在原点的力" 作用下的运动. 就是说, 我们的构形空间 M 被取为 $\mathbb{E}^3-\{0\}$, 而黎曼度量是 $m\times$ 欧几里得度量, 其中 m 为此质点的 "质量". 我们还假定函数 $\overline{U}$ 只依赖于到原点的距离. 因而 $\overline{U}(x,y,z)=P(r), r^2=x^2+y^2+z^2$. 在这些条件下可清楚看出, 绕原点的任一旋转是此黎曼度量下的等距变换, 并保持 $\overline{U}$ 不变. 因此我们可以把命题 5.2 应用于无穷小旋转 $x(\partial/\partial y)-y(\partial/\partial x)$, 从而得出对应的动量函数 xp_y-yp_x 为常值.(这个动量以绕 z 轴的角动量而得名). 相似地, 函数 xp_z-zp_x 和 yp_z-zp_y 也必在这个流的轨线上为常值. 如果记 $\boldsymbol{x}=\prec x,y,z\succ, \boldsymbol{p}=\prec p_x,p_y,p_z\succ$, 则这三个守恒定律可以联起来读作

$$\boldsymbol{x}\wedge\boldsymbol{p}=\text{常数}. \tag{6.1}$$

它被称为*角动量守恒定律*. 这里的 $\boldsymbol{x}$ 和 $\boldsymbol{p}$ 被看成 $T^*(M)$ 上的向量值函数. 为了研究 (6.1) 式所包含的意义, 我们先区分两种情形: 其常数 $\neq 0$ 的和为零的. 如果 (6.1) 式中出现的常数 $\neq 0$, 则 (6.1) 式表明由 $\boldsymbol{x}$ 和 $\boldsymbol{p}$ 张成的平面不变. 特别, 这个运动是使 $\boldsymbol{x}$ 保持在某个固定的平面中. 如果 $\boldsymbol{x}\wedge\boldsymbol{p}=0$, 我们讨论如下:

由于 $\boldsymbol{p}=m\dot{\boldsymbol{x}}$, 故 $\boldsymbol{x}\wedge\dot{\boldsymbol{x}}=0$. 因 $\boldsymbol{x}\neq 0$, 这表明

$$\dot{\boldsymbol{x}}=\lambda\boldsymbol{x}$$

对某个时间的函数 λ 成立. 现在 $\|\boldsymbol{x}\|=(\boldsymbol{x},\boldsymbol{x})^{1/2}$, 故有

$$\frac{d}{dt}\|\boldsymbol{x}\|=\frac{1}{\|\boldsymbol{x}\|}(\boldsymbol{x},\dot{\boldsymbol{x}})=\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|},\dot{\boldsymbol{x}}\right)=\lambda\|\boldsymbol{x}\| \tag{6.2}$$

因此

$$\frac{d}{dt}\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right)=\frac{\lambda\boldsymbol{x}}{\|\boldsymbol{x}\|}-\frac{\lambda\|\boldsymbol{x}\|\boldsymbol{x}}{\|\boldsymbol{x}\|^2}=0,$$

从而 $\boldsymbol{x}/\|\boldsymbol{x}\|=$ 常值向量; 即 x 在一条通过原点的固定射线上.

如果微分 (6.2) 式并利用 $\dfrac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ 为常数这个事实, 我们则由 (5.6) 式得到

$$m\frac{d^2}{dt^2}\|\boldsymbol{x}\|=\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|},m\frac{d}{dt}\dot{\boldsymbol{x}}\right)=-\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|},P'(\|\boldsymbol{x}\|)\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\right),$$

这就是

$$m\frac{d^2r}{dt^2}=-P'(r), \tag{6.3}$$

其中 $r = \|\boldsymbol{x}\|$.

不管怎样, 在所有情形中质点 $\boldsymbol{x}$ 总在一个平面中运动. 我们因此可集中注意到平面上. 就是说, 我们可以考虑一个 "新的" 力学系统, 那里的 $\overline{M} = \mathbb{E}^2 - \{0\}$, 以及它的黎曼度量, 函数 $\overline{U} = P(r)$, 等等. 让我们在这个平面中引进极坐标. 于是 $\partial/\partial\theta$ 保持了度量和位势不变. 如果 $\prec r, \theta, p_r, p_\theta \succ$ 为 $T^*(\overline{M})$ 上的对应坐标, 那么命题 5.2 表明

$$p_\theta = 常数 \tag{6.4}$$

(注意这确实是 (6.1) 式中我们不曾用过的部分.) 现在在极坐标下, 欧氏度量具有第九章的等式 (8.7) 的形式, 从而与质量 m 相关联的 $m\times$ 欧氏度量的黎曼度量由

$$\|(\dot{r}, \dot{\theta})\|^2 = m\{\dot{r}^2 + r^2\dot{\theta}^2\}.$$

给出. 特别, 从 $T(M)$ 到 $T^*(M)$ 的相关联的映射由

$$p_r = m\dot{r}, p_\theta = mr^2\dot{\theta} \tag{6.5}$$

给出. 因此 (6.4) 式表明

$$r^2\dot{\theta} = 常数. \tag{6.4$'$}$$

要了解 (6.4′) 式的意义, 我们考虑此平面中一条曲线 $\boldsymbol{x}(\cdot)$. 考虑由 0 到 $\boldsymbol{x}(0)$ 的射线的部分, 从 0 到 $\boldsymbol{x}(t)$ 的射线部分和从 0 到 t 的曲线 $\boldsymbol{x}(\cdot)$ 的部分围成的区域 u(见图 13.1), 于是

$$\frac{1}{2}\int_{\partial u} r^2 d\theta = \int_u r dr \wedge d\theta$$

是 u 的面积. 另一方面, 因为在射线上 $d\theta = 0$, 我们看到

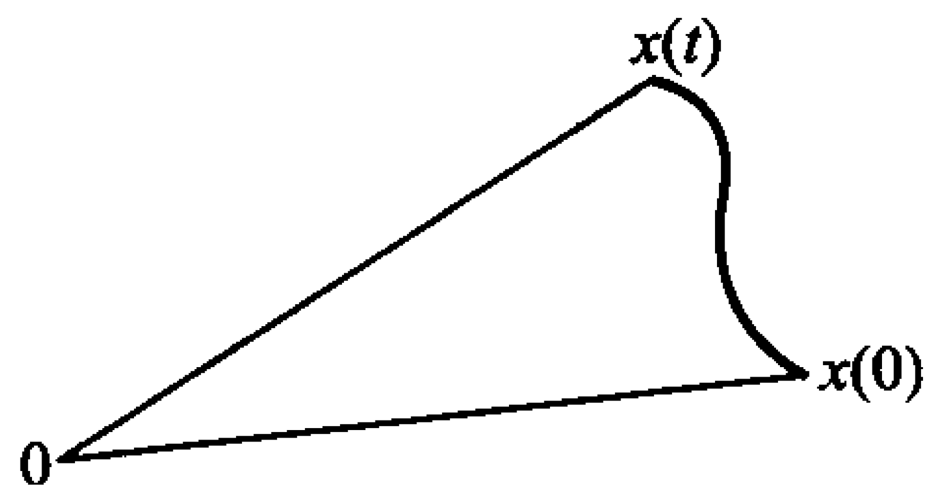


图 13.1

$$\int_{\partial u} r^2 d\theta = \int_0^t r^2\dot{\theta} dt.$$

因此 (6.4′) 式正是开普勒第二定律的内容: *质点以常速率扫过面积.*

因此我们已经看到, 系统的球面对称性表明这个运动在一个平面中并满足开普勒第二定律.

我们还没有利用到能量守恒, 在目前的情形下它可以读成

$$\frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + P(r) = 常数 = \frac{1}{2m}p_r^2 + \frac{1}{2mr^2}p_\theta^2 + P(r). \tag{6.6}$$

我们来考查满足 $p_\theta = A =$ 常数的一个特定的解曲线. 那么, 微分 (6.6) 式给出

$$\frac{dp_r}{dt} = m\frac{d^2r}{dt^2} = \frac{A^2}{mr^3} - P'(r). \tag{6.7}$$

我们可以把 (6.7) 式解释为一个一维力学系统的运动方程. 此系统的动能为 $\frac{1}{2}m\dot{r}^2$, 而势能 Q 由

$$Q(r) = P(r) + \frac{A^2}{2mr^2} \tag{6.8}$$

给出, (6.8) 式中的第二项被叫做向心位势, 而在 (6.7) 式中出现的对应项 A^2/mr^3 被叫做向心力.注意, 如果 $A = 0$, 则 (6.7) 式化为 (6.3) 式, 这正是我们所预期的.

现在我们能够用下面的步骤求解运动方程: 首先, 对所给的角动量求 (6.7) 式的各种解曲线. 对一个解 $r(\cdot)$, 通过积分 $\dot{\theta} = A/mr^2$ 决定 $\theta(\cdot)$, 得到

$$\theta(t) = \int_{t_0}^{t} \frac{A}{mr^2(s)} ds + \text{常数}. \tag{6.9}$$

我们运用能量守恒定律可以得到关于 (6.7) 式解的性质相当多的信息. 画一条像图 13.2 中那样的曲线. 假设轨线 $r(\cdot)$ 具有常值能量 $E_1 = \frac{1}{2}m\dot{r}^2 + Q(r)$, 于是如果集合 $\{r : Q(r) \leqslant E_1\}$ 有界, 并且因为 $\frac{1}{2}m\dot{r}^2 \geqslant 0$, 从而 $r(t)$ 必定总在此集合中. 事实上, 假设区间 $a \leqslant r \leqslant b$ 使得 $Q(a) = Q(b) = E_1$, 并且对小的 ε 有 $Q(b+\varepsilon) > E_1, Q(a-\varepsilon) > E_1$. 于是, 如果 $a \leqslant r(t_0) \leqslant b$ 对某个值 t_0 成立, 则得到对所有 t 有 $a \leqslant r(t) \leqslant b$. 另外, 如果 $Q(r) < E_1$ 对 $a < r < b$ 成立, 则对于 $a < r(t_0) < b$, $\dot{r}(t)$ 不能为零. 从而这个质点移向极限 $r = a$ 和 $r = b$ 中的一个. 如果我们用能量守恒定律, 我们看到

$$\frac{1}{2}m\dot{r}^2 + Q(r) = E_1,$$

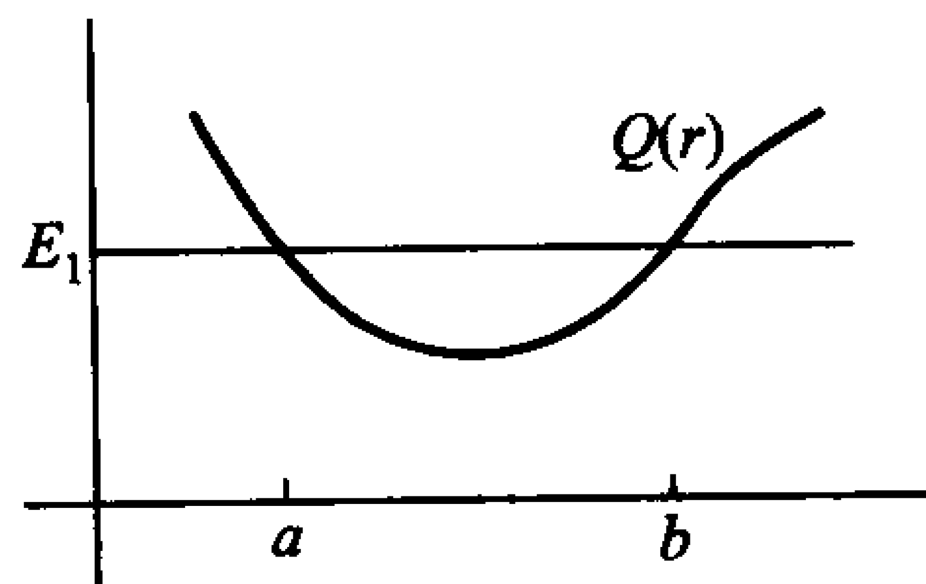


图 13.2

所以

$$\frac{dr}{dt} = \dot{r} = \pm\sqrt{\frac{E_1 - Q(r)}{\frac{1}{2}m}}.$$

对 $a < r < b$ 我们看出 r 是 t 的一个单调函数, 我们可以通过公式

$$t(r) = \pm\left(\frac{1}{2}m\right)^{1/2}\int_{t_0}^{r}\frac{dx}{\sqrt{F_1 - Q(x)}} + t_0. \tag{6.10}$$

得到作为 r 函数的 t, 其中 $r(t_0) = r_0$. 特别地, 如果 $Q'(a) \neq 0, Q'(b) \neq 0$, 则 (6.10) 式中的积分收敛, 从而对 r 可取有限的时间值得到 a 或 b. 因此, (因为 $d\dot{r}/dt = Q'(r) \neq 0$) 函数 r 在 a 和 b 之间摆动, 从一头到另一头所花时间为

$$T = \left(\frac{1}{2}m\right)^{1/2}\int_{a}^{b}\frac{dx}{\sqrt{E_1 - Q(x)}}.$$

如果 $\{r : Q(r) \leqslant E_2\}$ 无界, 则具能量 E_2 的运动也不必有界. 例如, 假设对 $r > a, Q(r) < E_2$. 于是当 $r(t_0) > a$, 并且 $\dot{r}(t_0) < 0$ 时, r 将减少直至达到 a 的那个时刻便掉回头, 然后消失到无穷远; 如果 $\dot{r}(t_0) > 0$, 则 r 径直走向无穷. 这时的轨线是 r 从无穷处进入, 到 a 处掉头, 然后又走回到无穷的轨线.

这使我们想起了函数 $r(\cdot)$ 的描述性的分析. 曲线 $\theta(\cdot)$ 于是被 (6.9) 式所决定.

我们常常感兴趣的轨线是在 $r\theta$ 平面中与时间依赖性无关的曲线. 假设能量为 E, 角动量为 A . 对于 $r \neq 0$ 的范围, 我们可将

$$\frac{d\theta}{dt} = \frac{A}{mr^2}$$

代入 (6.10) 式得到 [由于 $d\theta/dr = (d\theta/dt)(dt/dr)$]

$$\theta(r) = \pm\left(\frac{1}{2m}\right)^{1/2}\int_{r_0}^{r}\frac{Adx}{x^2\sqrt{E - Q(x)}} + \theta(r_0).$$

对应于万有引力的平方反比律的 $P(r) = -\alpha/r$ 这个重要情形中, 我们有

$$Q(x) = \frac{A^2}{2mx^2} - \frac{\alpha}{x},$$

所以被积项为

$$\frac{1}{2(m)^{1/2}}\cdot\frac{A}{x^2\sqrt{E-\dfrac{A^2}{2mx^2}+\dfrac{\alpha}{x}}}=\frac{A}{x^2\left[2mE-\left(\dfrac{A}{x}-\dfrac{m\alpha}{A}\right)^2+\dfrac{m^2\alpha^2}{A^2}\right]^{1/2}}$$

$$=\frac{d}{dx}\left[\arccos\frac{\dfrac{A}{x}-\dfrac{m\alpha}{A}}{\sqrt{2mE+\dfrac{m^2\alpha^2}{A^2}}}\right].$$

从而

$$\theta(r)-\theta(r_0)=\arccos\frac{\dfrac{A}{r}-\dfrac{m\alpha}{A}}{\sqrt{2mE+\dfrac{m^2\alpha^2}{A^2}}}.$$

我们选取 r_0 为 r 的最小值, 并选取 $\theta(r_0)=0$. 同时令

$$p=\frac{A^2}{m\alpha},e=\sqrt{1+\frac{1EA^2}{m\alpha^2}}.$$

我们看到轨道的方程是

$$\frac{p}{r}=1+e\cos\theta. \tag{6.11}$$

这是条圆锥截线的方程 (开普勒第一定律).如果 $E<0$, 则 $e<1$, 那么 (6.11) 式代表了一个椭圆, 其长短半轴分别为

$$a=\frac{p}{1-e^2}=\frac{\alpha}{2|E|}\quad 和\quad b=\frac{p}{\sqrt{1-e^2}}=\frac{A}{\sqrt{2m|E|}}.$$

我们把做出双曲线和抛物线轨道的细节留给读者.

对于椭圆轨道, 开普勒第二定律表明椭圆内部的总面积按匀速 $A/2m$ 被扫过. 因此椭圆的面积 πab 由 $(A/2m)T$ 给出. 从而

$$\frac{A}{2m}T=\pi ab=\pi a\frac{A}{\sqrt{2m|E|}},$$

由此得到

$$T=2\pi a^{3/2}\sqrt{m/\alpha}.$$

换而言之, 运动周期的平方与轨道的线性量纲成的立方比例 (开普勒第三定律).

13.7 二体问题

我们考虑由 n 个质点组成的力学系统, 其中每个点具质量 m_i, 从而构形空间为 $\mathbb{E}^3 \times \cdots \times \mathbb{E}^3$ (n 次), 并且, 如果 $a = \prec a^1, \cdots, a^n \succ \in M$, 其中 $a^i = \prec x^i, y^i, z^i \succ \in \mathbb{E}^3$, 于是

$$\begin{aligned}\frac{1}{2}\|\dot{a}\|^2 &= \frac{1}{2}m_1\|\dot{a}^1\|^2 + \cdots + \frac{1}{2}m_n\|\dot{a}^n\|^2 \\ &= \frac{1}{2}m_1[(\dot{x}^1)^2 + (\dot{y}^1)^2 + (\dot{z}^1)^2] + \cdots + \frac{1}{2}m_n[(\dot{x}^n)^2 + (\dot{y}^n)^2 + (\dot{z}^n)^2]\end{aligned}$$

为此系统的动能. 与我们在第 1 节表示的一样, 我们希望限制 M 为 $\mathbb{E}^3 \times \cdots \times \mathbb{E}^3$ 中那样的子集, 使得对任意 $i \neq j, a^i \neq a^j$. 我们还进一步假设势能只依赖于质点间的相互距离. 就是说, 假定

$$\overline{U}(a^1, \cdots, a^n) = P(\|a^2 - a^1\|, \cdots, \|a^j - a^1\|, \cdots, \|a^n - a^{n-1}\|).$$

那么, 如果 A 是 $\mathbb{E}^3$ 的欧几里得运动并将 A 同时用于所有的质点, 则动能和势能守恒. 即变换

$$\prec a^1, \cdots, a^n \succ \mapsto \prec Aa^1, \cdots, Aa^n \succ$$

是 M 上黎曼度量的等距映射并保持 $\overline{U}$ 不变.

设 $\prec x^1, y^1, z^1, \cdots, x^n, y^n, z^n \succ$ 为 M 上的笛卡儿坐标, 并设

$$\prec q_{x^1}, q_{y^1}, q_{z^1}, q_{x^2}, \cdots, q_{y^n}, q_{z^n}, p_{x^1}, \cdots, p_{z^n} \succ$$

为 $T^*(M)$ 上的相应坐标. 现在 $\partial/\partial x$ 代表 $\mathbb{E}^3$ 中在 x 方向上无穷小平移的向量场. 因此,

$$\frac{\partial}{\partial x^1} + \frac{\partial}{\partial x^2} + \cdots + \frac{\partial}{\partial x^n}$$

是单参数群

$$\prec x^1, y^1, z^1, \cdots, x^n, y^n, z^n \succ \mapsto \prec x^1 + t, y^1, z^1, x^2 + t, y^2, z^2, \cdots, x^n + t, y^n, z^n \succ$$

的无穷小生成元. 我们因此可应用命题 5.2 得出结论说函数

$$p_{x^1} + p_{x^2} + \cdots + p_{x^n}$$

在任一轨线上为常值. 这个函数以 x 方向的总线性动量而知名. 相似地, y 方向的总线性动量为

$$p_{y^1} + p_{y^2} + \cdots + p_{y^n},$$

而 z 方向的总线性动量便是

$$p_{z^1}+p_{z^2}+\cdots+p_{z^n},$$

它们必定守恒. 如果我们定义 $p^i\in\mathbb{E}^3$ 为 $p^i=\prec p_{x^i},p_{y^i},p_{z^i}\succ$, 我们则可以说 $\mathbb{E}^3$ - 值函数

$$p^1+p^2+\cdots+p^n$$

必定守恒. 这是总线性动量守恒定律.(对于两个质点的情形断言 p_1+p_2 守恒恰好是牛顿的 "作用与反作用相等" 的定律. 在我们的架构下, 我们看到的这个定律是在平移下物理状态不变性的一种反映.)

向量场 $x(\partial/\partial y)-y(\partial/\partial x)$ 代表了 $\mathbb{E}^3$ 中绕 z 轴的无穷小旋转. 因此

$$x^1\frac{\partial}{\partial y^1}-y^1\frac{\partial}{\partial x^1}+x^2\frac{\partial}{\partial y^2}-y^2\frac{\partial}{\partial x^2}+\cdots+x^n\frac{\partial}{\partial y^n}-y^n\frac{\partial}{\partial x^n}$$

代表了所有质点同时绕 z 轴的无穷小旋转. 因而函数

$$q_{x^1}p_{y^1}-q_{y^1}p_{x^1}+\cdots+q_{x^n}p_{y^n}-q_{y^n}p_{x^n}$$

守恒. 这是绕 z 轴的总角动量守恒定律. 相似地, 我们得到绕 x 和 y 轴的总角动量守恒定律. 如果我们令

$$q^i=\prec q_{x^i},q_{y^i},q_{z^i}\succ\in\mathbb{E}^3,$$

我们则可以把这三个方程联合起来说在 $\wedge^2(\mathbb{E}^3)$ 取值的函数

$$q^1\wedge p^1+\cdots+q^n\wedge p^n$$

在运动轨线上为常值.

让我们更仔细地考查总线性动量守恒定律. 由于有

$$p_{x^i}=m_i\frac{dx^i}{dt}$$

等等这类事实, 我们便有 $p^i=m_i(da^i/dt)$, 从而

$$\frac{d}{dt}(m_1a^1+m_2a^2+\cdots+m_na^n)=\text{常数}.$$

换句话说, 质心

$$C=\frac{m_1a^1+\cdots+m_na^n}{m_1+\cdots+m_n}$$

以不变速度沿直线运动.

假设只有两个质点. 那么合理的做法是引进质心

$$C = \frac{m_1 a^1 + m_2 a^2}{m_1 + m_2}$$

和相对位置向量

$$d = a^1 - a^2$$

作为新的坐标. 如果解出 a^1, a^2, 得到

$$a^1 = \frac{m_2 d}{m_1 + m_2} + C, a^2 = -\frac{m_1 d}{m_1 + m_2} + C,$$

因此动能由

$$K(C, d) = \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}\|d\|^2 + \frac{1}{2}(m_1 + m_2)\|C\|^2$$

给出, 而 $\overline{U}(c, d) = P(\|d\|)$.

那么, 这个系统的运动有下面的特征描述: 质心作匀速直线运动, 而相对位置向量 $d = a^1 - a^2$ 满足具质量

$$\frac{m_1 m_2}{m_1 + m_2}$$

的单个质点的运动方程, 此质点位于中心力场, 具势能 $U = P(\|d\|)$. 因而可用前面一节结果来确定此质点的运动. 特别, 如果 $P(\|d\|) = \alpha/\|d\|$ 是反平方比势能, 则对应的二体问题能够被完全解决.

注意, 如果 m_2 相对于 m_1 而言非常之大, 则 C 非常接近于 a^2. 这时的 $d(\cdot)$ 是对相对于大的一个质点而言的那个较小质点运动的一个好的逼近.

这种情形例如可发生在研究相对于太阳的行星运动中.

13.8 拉格朗日方程

我们的力学讨论把我们引向到在 $T^*(M)$ 上的某类向量场 X, 其中 M 是构形空间. 在 $H = K + U$ 的情形, 给出 K 的黎曼度量诱导了一个从 $T(M)$ 到 $T^*(M)$ 的同构 $\mathfrak{L}$(即一个微分同胚, 它在每个 $T_x(M)$ 上是线性的). 我们因此得到 $T(M)$ 上一个向量场 Y, 使得 $\mathfrak{L}_* Y_v = X_{\mathfrak{L}(v)}, v \in T(M)$ 为任意. 于是我们可以问

及关于这个向量场 Y 在 $T(M)$ 的与 M 的坐标 $\prec x^1,\cdots,x^n \succ$ 相关联的局部坐标 $\prec q^1,\cdots,q^n,\dot{q}^1,\cdots,\dot{q}^n \succ$ 下的形状是什么的问题. 假设其黎曼度量由

$$\|(\dot{q}^1,\cdots,\dot{q}^n)\|^2 = \sum g_{ij}(q^1,\cdots,q^n)\dot{q}^i\dot{q}^j$$

给出. 如果 $\prec q^1,\cdots,q^n,p^1,\cdots,p^n \succ$ 为 $T^*(M)$ 上对应的局部坐标, 则微分同胚 $\mathfrak{L}$ 便由

$$q^i = q^i, p^i = \sum g_{ij}(q^1,\cdots,q^n)\dot{q}^j \tag{8.1}$$

给出.

我们能够应用 (8.1) 式去得出 $\mathfrak{L}$ 的局部表达式, 从而 Y 的局部表达式. 然而更方便的是作一点不同的论证. 设 L 为 $T(M)$ 上的一个函数, 定义为

$$L(q^1,\cdots,q^n,\dot{q}^1,\cdots,\dot{q}^n) = \frac{1}{2}\sum g_{ij}\dot{q}^i\dot{q}^j - \overline{U}(q^1,\cdots,q^n), \tag{8.2}$$

即 $L = K - U$, 其中 $K(v) = \frac{1}{2}\|v\|^2, U(v) = \overline{U}(\pi(v))$ 分别为表示成 $T(M)$ 上函数的动能和势能. 我们可重写 (8.1) 式为

$$q^i = q^i, p^i = \frac{\partial L}{\partial \dot{q}^i}. \tag{8.3}$$

于是对任意 $l \in T^*(M)$ 我们有

$$H(l) = \langle \mathfrak{L}^{-1}(l), l\rangle - L(\mathfrak{L}^{-1}(l)). \tag{8.4}$$

事实上, 由定义,

$$\langle \mathfrak{L}^{-1}(l), l\rangle = \|l\|^2 = \|\mathfrak{L}^{-1}(l)\|^2,$$

从而

$$\begin{aligned}\langle \mathfrak{L}^{-1}(l), l\rangle - L(\mathfrak{L}^{-1}(l)) &= \|l\|^2 - \frac{1}{2}\|l\|^2 + \overline{U}(\pi(l)) \\ &= \frac{1}{2}\|l\|^2 + U(l) = H(l).\end{aligned}$$

依据局部坐标我们可以把 (8.4) 式写为

$$H(q^1,\cdots,q^n,p^1,\cdots,p^n) = \sum p^i\dot{q}^j - L(q^1,\cdots,q^n,\dot{q}^1,\cdots,\dot{q}^n), \tag{8.5}$$

其中的 $\dot{q}^i$ 通过 $\mathfrak{L}^{-1}$ 看成是这些 q 和 p 的函数. 由于 $H(l)=\frac{1}{2}||l||^2+\overline{U}(\pi(l))$, 我们可将映射 $\mathfrak{L}^{-1}$ 写成

$$q^i=q^i,\dot{q}^i=\frac{\partial H}{\partial p^i}. \tag{8.6}$$

又由 (8.5) 式有

$$\begin{aligned}\frac{\partial H}{\partial q^i}&=\frac{\partial}{\partial q^i}\left[\sum\dot{q}^ip^i-L(q^1,\cdots,q^n,\dot{q}^1,\cdots,\dot{q}^n)\right]\\&=\sum\frac{\partial(\dot{q}^i\circ\mathfrak{L}^{-1})}{\partial q^i}p^i-\frac{\partial(L\circ\mathfrak{L}^{-1})}{\partial q^i}-\frac{\partial L}{\partial q^i}\frac{\partial\dot{q}^i\circ\mathfrak{L}^{-1}}{\partial q^i}\\&=-\frac{\partial L\circ\mathfrak{L}^{-1}}{\partial q^i},\end{aligned}$$

其中用了 (8.4) 式.

现在设 $v(\cdot)=\prec q^1(\cdot),\cdots,q^n(\cdot),\dot{q}^1(\cdot),\cdots,\dot{q}^n(\cdot)\succ$ 为向量场 Y 的一条解曲线, 故 $\mathfrak{L}\circ v(\cdot)$ 是 $T^*(M)$ 上 X 的一条解曲线. 于是由 (5.1) 式,

$$\frac{dq^i}{dt}=\frac{\partial H}{\partial p^i}=\dot{q}^i$$

以及

$$\frac{dp^i}{dt}=\frac{d(\partial L/\partial\dot{q}^i)}{dt}=-\frac{\partial H}{\partial q^i}=\frac{\partial L}{\partial q^i}.$$

换句话说, 方程 (5.1) 等价于在 $T(M)$ 上的方程

$$\frac{dq^i}{dt}=\dot{q}^i,\quad\frac{d(\partial L/\partial\dot{q}^i)}{dt}-\frac{\partial L}{\partial q^i}=0. \tag{8.7}$$

方程 (8.7) 的第一个是说我们现在所处理的是个二阶微分方程组, 它由方程 (8.7) 中的第二个隐式给出. 称方程 (8.7) 为拉格朗日方程.

对某些目标而言拉格朗日方程非常之方便. 我们用建立“力学相似原理”来解释这点. 注意, 方程 (8.7) 在将 L 换成 cL 时保持不变, 其中的 c 为任意非零常数. 假设 M 是某个具线性坐标 $x^1,\cdots,x^n$ 的向量空间中的一个开子集, 而其上的黎曼度量由 $\sum g_{ij}\dot{q}^i\dot{q}^j$ 给出, 其中的 g_{ij} 为常数. 设 m_α 是由乘以 $\alpha>0$ 构成的线性变换, 即 $m_\alpha\prec x^1,\cdots,x^n\succ=\prec\alpha x^1,\cdots,\alpha x^n\succ$. 假设 $\overline{U}$ 是个 p 次齐次函数, 使得

$$\overline{U}(\alpha x^1,\cdots,\alpha x^n)=\alpha^p\overline{U}(x^1,\cdots,x^n).$$

现在我们以一个因子 β 改变时间刻度, 即以 $s=\beta t$ 代替 t. 于是

$$\frac{d\alpha q^i}{ds}=\frac{\alpha}{\beta}\frac{dq^i}{dt},$$

且

$$\frac{1}{2}\sum g_{ij}\frac{d\alpha q^i}{ds}\frac{d\alpha q^j}{ds}=\frac{\alpha^2}{\beta^2}\frac{1}{2}\sum g_{ij}\frac{dq^i}{dt}\frac{dq^j}{dt}.$$

选取 β 使 $\alpha^2/\beta^2=\alpha^p$, 即 $\beta=\alpha^{1-(1/2)p}$. 于是

$$L\left(\alpha q^1,\cdots,\alpha q^n,\frac{\alpha dq^1}{ds},\cdots,\frac{\alpha dq^n}{ds}\right)=\alpha^p L\left(q^1,\cdots,q^n,\frac{dq^1}{dt},\cdots,\frac{dq^n}{dt}\right).$$

换句话说, 将 q^i 换为 αq^i, t 换为 βt 则把解变成了解: 如果以 α 改变线性刻度, 以 $\alpha^{1-(1/2)p}$ 改变时间刻度, 我们则得一个同构的情形. 例如, 如果 U 是 -1 次齐次的 (如同在引力的平方反比定律中那样), 则 $\beta=\alpha^{3/2}$. 特别, 任何一个周期轨道的周期与它的线性量纲的 3/2 幂成比例, 这正是开普勒第三定律. 因而我们看出开普勒第三定律完全是引力的平方反比定律的推论.

回到一般拉格朗日系统的讨论, 我们再次以它是个二阶微分方程组的角度观察它. 因此我们可以对它应用基本存在性和惟一性定理得出结论说, 对每个 $x\in M$ 和每个 $\xi\in T_x(M)$, 存在惟一的曲线 $C(\cdot)$, 它是此系统对 $C(0)=x, C'(0)=\xi$ 的一条轨线.

13.9 变分原理

函数 L 在力学的变分原理的研究中起着关键作用. 考虑下面的问题: 设 p 和 q 是 M 中二个点, $t_1<t_2$ 为两个实数. 对定义在区间 $[t_1,t_2]$ 上的任意一条可微曲线 C 我们令

$$I[C]=\int_{t_1}^{t_2}L(C'(t))dt. \tag{9.1}$$

这里 $C'(t)\in T(M)$ 而 L 是 $T(M)$ 上的函数, 故此积分有意义. 在所有联结 p 到 q 的曲线中我们想要找出使 $I[C]$ 取最小值的那条曲线. 我们将看到使 $I[C]$ 最小的一个必要条件是它是某个力学系统的轨线.(事实上, 如果在曲线构成的空间中引进邻域的适当概念, 那么它也是 C 甚至为局部极小的一个必要条件.)

在建立这个结果之前, 给出 $I[C]$ 另一个使用起来方便的表达式. 像前面那样, 设 $\mathfrak{L}$ 为由黎曼度量给出的映射 $T(M)\to T^*(M)$. 设 $\overline{C}$ 为 $T^*(M)$ 中由

$\overline{C}(t) = \mathfrak{L}(C'(t))$ 给出的曲线. 于是,

$$I[C] = \int_{\overline{C}} \theta - \int_{t_1}^{t_2} H(\overline{C}(t))dt, \tag{9.2}$$

其中 θ 为 $T^*(M)$ 上的基本线性形式.

事实上, 利用局部坐标,

$$\int_{\overline{C}} \theta = \sum \int_{\overline{C}} p^i dq^i = \sum \int_{C'} \frac{\partial L}{\partial \dot{q}^i} dq^i = \sum \int \frac{\partial L}{\partial \dot{q}^i} \dot{q}^i(t)dt,$$

这是由于由定义, C' 是使得

$$\dot{q}^i(t) = \frac{dq^i}{dt}(t)$$

的曲线. 但是由 (8.5) 式,

$$\sum \int \frac{\partial L}{\partial q^i} \dot{q}^i(t)dt = \sum \int (p^i \circ \mathfrak{L})(t)\dot{q}^i(t)dt = \int [H \circ \mathfrak{L}(C'(t)) + L(C'(t))]dt,$$

故 (9.2) 式成立.

设 Z 为 M 上的向量场, 它生成一个流 φ. 对所有充分小的 s, 曲线 $\varphi_s \circ C(\cdot)$ 有定义, 并且

$$(\varphi_s \circ C)'(t) = \varphi_{s*}C'(t) = T[\varphi_s](C'(t)),$$

所以

$$\begin{aligned} I[\varphi_s \circ C] &= \int_{t_1}^{t_2} L(T[\varphi_s]C'(t))dt \\ &= \int_{\mathfrak{L}\circ\varphi_s\circ C'} \theta - \int_{t_1}^{t_2} H(\mathfrak{L} \circ T[\varphi_s](C'(t)))dt \end{aligned} \tag{9.3}$$

由于 $I[C]$ 是个最小值, 我们必有

$$\frac{dI(\varphi_s \circ C)}{ds} = 0.$$

我们现在要计算这个导数以导出上面方程的一个结果.

现在 $\mathfrak{L} \circ T[\varphi_s] \circ \mathfrak{L}^{-1}$ 是 $T^*(M)$ 上的一个流, 它满足

$$\pi \circ \mathfrak{L} \circ T[\varphi_s] \circ \mathfrak{L}^{-1} = \varphi_s.$$

设 $\overline{Z}$ 为这个流的无穷小生成元, 使得在 $T^*(M)$ 上的所有点有

$$\pi_*\overline{Z} = Z. \tag{9.4}$$

如果在 $s=0$ 作关于 s 的 (9.3) 式的微分, 我们得到

$$\frac{d}{ds}I[\varphi_s \circ C] = \int_{\overline{C}} D_{\overline{Z}}\theta - \int_{t_1}^{t_2} D_{\overline{Z}}H(\overline{C}(t))dt.$$

现在

$$\begin{aligned} D_{\overline{Z}}\theta &= d\langle \overline{Z}, \theta\rangle + \overline{Z}\rfloor d\theta, \langle \overline{Z}, \theta\rangle_l = \langle \pi_*\overline{Z}, l\rangle \\ &= \langle Z, l\rangle = f_Z(l), \end{aligned}$$

所以得到

$$\frac{d}{ds}I[\varphi_s \circ C] = \int_{\overline{C}} \overline{Z}\rfloor d\theta - \int_{t_1}^{t_2} D_{\overline{Z}}H(\overline{C}(t))dt + f_Z(\overline{C}(t_2)) - f_Z(\overline{C}(t_1)). \tag{9.5}$$

现假设向量场 Z 在 p 和 q 为零. 于是所有的曲线 $\varphi_s \circ C$ 都联结了 p 到 q. 如果 C 使积分 I 极小, 则导数 $d(I[\varphi_s \circ C])/ds$ 在 $s=0$ 必为 0, 注意, 这时 (9.5) 式的最后两项为零, 因而对所有在 p 和 q 为 0 的向量场必有

$$\int_{\overline{C}} \overline{Z}\rfloor d\theta - \int_{t_1}^{t_2} D_Z H(\overline{C}(t)) = 0 \tag{9.6}$$

特别地, 我们取

$$\begin{aligned} Z &= \psi\frac{\partial}{\partial x^i}, \text{ 当} x \in U, \\ &= 0, \text{ 当} x \notin U, \end{aligned}$$

其中 ψ 为一个 C^∞ 函数, 其支集在 M 的一个坐标邻域中而此邻域的坐标为 $\prec x^1, \cdots, x^n \succ$. 假设进而有 $\psi(p) = \psi(q) = 0$, 其中 $p \in U$ 或者 $q \in U$. 那么由 (2.7′) 式知

$$T(Z) = \psi\frac{\partial}{\partial q^i} + \sum \frac{\partial \psi}{\partial q^j}\frac{\partial}{\partial \dot{q}^j},$$

故 $\overline{Z} = \mathfrak{L}_*T(Z) = \psi\frac{\partial}{\partial q^i} + \sum B^j \frac{\partial}{\partial p^j}$,

这里的 B^j 为 $T^*(M)$ 上的某一函数, 它线性地依赖于 Z.[当然, 这只是 (9.4) 式的一个重新陈述.] 于是

$$\overline{Z}\rfloor d\theta = \overline{Z}\rfloor \sum dp^i \wedge dq^i = \sum_j B^j dq^j - \psi dp^i,$$

同时

$$D_{\overline{Z}}H = \psi\frac{\partial H}{\partial q^i} + \sum_j B^j\frac{\partial H}{\partial p^j}.$$

因此, 如果曲线 $\overline{C}$ 由 $\overline{C}(t) = \prec q^1(t),\cdots,q^n(t),p^1(t),\cdots,p^n(t)\succ$ 给出, 那么方程 (9.6) 成为

$$\int\left[\sum B^j\left(\frac{dq^j}{dt}-\frac{\partial H}{\partial p^j}\right)-\psi\left(\frac{dp^i}{dt}+\frac{\partial H}{\partial q^i}\right)\right]dt = 0$$

现在由构造, $\overline{C}(t) = \mathfrak{L}\circ C'(t)$, 故在 $\overline{C}$ 上由 (8.6) 式我们有

$$\frac{dq^j}{dt} = \dot{q}^j = \frac{\partial H}{\partial p^j}. \tag{9.7}$$

在上面积分中的第一个和式必为 0, 因而我们必定有

$$\int\psi\left(\frac{dp^i}{dt}+\frac{\partial H}{\partial q^i}\right)dt = 0.$$

对于所有支集在一个坐标邻域中且在 p 和 q 为零的函数 ψ 它必定成立. 显然, 只有

$$\frac{dp^i}{dt} = -\frac{\partial H}{\partial q^i} \tag{9.8}$$

它才可能成立. 这对所有的 i 也必定成立. 因为 (9.6) 和 (9.8) 式完全就是 (5.2) 式, 我们从而可以断定:

命题 9.1　*C 使泛函 $I[C]$ 极小的一个必要条件是 C 为对应的动力系统的一条轨线.*

这个必要条件什么时候也是一个充分条件则是一个复杂的问题. 在这里我们对此不做任何一般性的讨论, 但建议读者参阅任何一本关于变分学的标准读本.

我们顺便从 (9.5) 式可以得到一个额外的结果. 假设我们考虑下述问题: 设 N_1 和 N_2 是 M 的两个子流形 (图 13.3), 并假设我们要求 C 在所有联结 N_1 和 N_2 的曲线中使 C 极小而不仅仅是在联结 p 到 q 的那些曲线中. 这时的 (9.5) 式不得不对所有在 p 切于 N_1 和在 q 切于 N_2 的向量场 Z 为 0. 现在我们看到, 如果 C 是这个极小问题的一个解, 它肯定是在所有联结 p 和 q 的曲线中极小 I 的问题的一个解. 特别, C 必是此力学系统的一条轨线. 读者容易验证出, 这意味着对所有向量场 Z 有

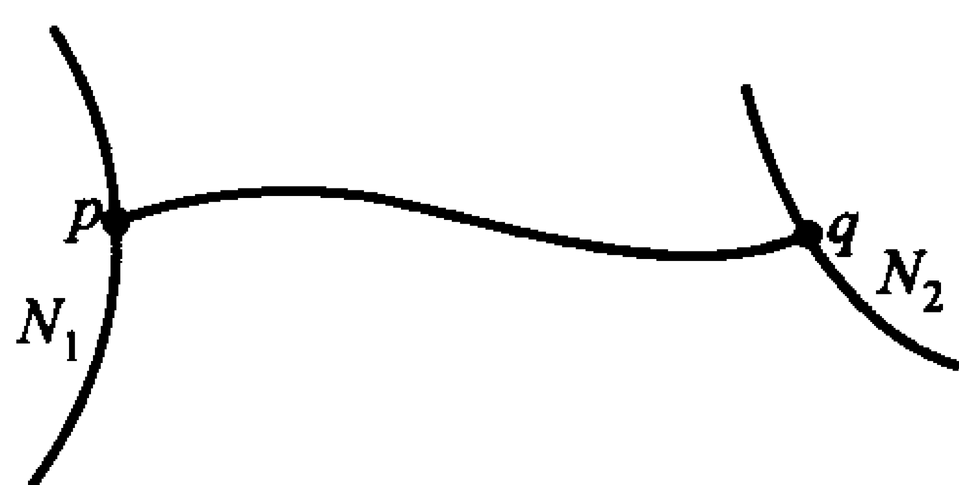


图 13.3

$$\int_{\overline{C}} \overline{Z} \rfloor \theta - \int (D_{\overline{Z}} H)(\overline{C}(t)) dt = 0$$

因此, 如果 C 是此更难的极小化问题的解, 我们则必有

$$f_Z(\overline{C}(t_1)) - f_Z(\overline{C}(t_2)) = 0.$$

现在

$$f_Z(\overline{C}(t_1)) = \langle Z_p, \overline{C}(t_1) \rangle, f_Z(\overline{C}(t_2)) = \langle Z_q, \overline{C}(t_2) \rangle.$$

由于当 $p \neq q$ 时我们可以在 $T_p(N_1)$ 中任意选择 Z_p, 在 $T_q(N_2)$ 中任意选取 Z_q, 故在此时我们得到结论说, C 必定也满足

$$\begin{aligned} &\langle \xi, \overline{C}(t_1) \rangle = 0, \quad \text{对所有} \xi \in T_p(N_1) \text{成立}, \\ &\langle \eta, \overline{C}(t_2) \rangle = 0, \quad \text{对所有} \eta \in T_q(N_2) \text{成立}. \end{aligned} \tag{9.9}$$

因为 $\overline{C}(t_1) = \mathfrak{L} C'(t_1)$, 我们有

$$\langle \xi, \overline{C}(t_1) \rangle = \langle \xi, \mathfrak{L} C'(t_1) \rangle = \langle \xi, C'(t_1) \rangle,$$

所以可将 (9.9) 式写成

$$\begin{aligned} &\langle \xi, C'(t_1) \rangle = 0, \quad \text{对所有} \xi \in T_p(N_1) \text{成立}, \\ &\langle \eta, C'(t_2) \rangle = 0, \quad \text{对所有} \eta \in T_q(N_2) \text{成立}. \end{aligned} \tag{9.10}$$

换句话说, 曲线 C 必定与子流形 N_1 和 N_2 正交.

虽然我们对命题 9.1 的叙述是在动力系统的框架下表达出来的, 但它实际上可以在一种更为一般的背景下加以阐述. 设 L 为 $T(M)$ 上的任一函数而不必是 $K-U$ 这种类型. 我们可以如在 (9.1) 式中那样地定义积分 I, 并且也提出极小化问题. 这是变分法中的典型问题. 在 3.15 节我们已经从不同的观点讨论过此

问题. 我们在矩阵 $\left(\frac{\partial L}{\partial \dot{x}^i \partial \dot{x}^j}\right)$ 无处奇异的条件下把我们的论证转移到这个更一般情形的任务留给了读者. (这里的映射 $\mathfrak{L}: T(M) \to T^*(M)$ 由

$$\prec x^1, \cdots, x^n, \dot{x}^1, \cdots, \dot{x}^n \succ \mapsto \prec x^1, \cdots, x^n, \frac{\partial L}{\partial \dot{x}^1}, \cdots, \frac{\partial L}{\partial \dot{x}^n} \succ,$$

从而上面的非异性保证了这个映射是个局部的微分同胚.)

13.10 测地坐标

这一节, 我们要暂时搁置力学研究以便展示前面几节的结果在黎曼流形研究中的一些应用. 我们注意到, 一个黎曼流形在令势能为零时总是由它的动能定义了一个力学系统. 这是一个特殊的力学系统类型; 我们在这一节中要研究它.

设 M 是具黎曼度量的一个有限维流形, 并如同前面那样, 定义 $T(M)$ 上的 L 为

$$L(v) = \frac{1}{2}\|v\|^2.$$

于是这定义了 $T(M)$ 上的一个向量场 Y, 它在局部坐标下对应于微分方程组 (8.7) . 设 p 为 M 中一点. 对每个 $\xi \in T_p(M)$ 存在一条惟一的轨线 $C_\xi(\cdot)$ 使得

$$C_\xi(0) = p, C'_\xi(0) = \xi.$$

利用局部坐标, 如果 $\xi = \prec \xi^1, \cdots, \xi^n \succ$, 则

$$C_\xi(t) = \prec q^1_\xi(t), \cdots, q^n_\xi(t) \succ,$$

其中 q^i_ξ 是微分方程

$$\frac{dq^i_\xi}{dt} = \dot{q}^i_\xi, \frac{d(\partial L/\partial \dot{q}^i)}{dt} - \frac{\partial L}{\partial q^i} = 0$$

在初始条件 $q^i_\xi(0) = x^i(p), \dot{q}^i_\xi(0) = \xi^i$ 下的惟一解. 由基本的存在性和惟一性定理知道, 函数 q^i_ξ 对充分小的 t 有定义. 我们可将 $C_\xi(t)$ 看为同时依于 ξ 和 t 的. 换句话说, 我们有一个映射 $C.(\cdot)$ 它对每个 $\xi \in T_p(M)$ 和每个充分小的 t 指定了 M 的一个点. 事实上映射 $C.(\cdot)$ 在 $T_p(M) \times \{0\} \subset T_p(M) \times \mathbb{R}$ 的某个邻域中有定义.

上面所说的对任何拉格朗日函数成立. 在势能为零的情形我们可以说得更多一些. 设 s 为一实数. 考虑曲线 $C_\xi(s\cdot)$, 即

$$C_\xi(st) = \prec q^1_\xi(st), \cdots, q^n_\xi(st) \succ .$$

于是 (隐去角标 ξ, 在下面计算中不说自明)

$$\frac{dq^i(st)}{dt} = s\dot{q}^i(st),$$

$$\frac{\partial L}{\partial q^i}(q^1(st),\cdots,q^n(st),s\dot{q}(st),\cdots,s\dot{q}^n(st)) = \frac{1}{2}s^2\sum\frac{\partial g_{kl}}{\partial q^i}\dot{q}^k(st)\dot{q}^l(st)$$

并且

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}^i}(q^1(st),\cdots,q^n(st),s\dot{q}^1(st),\cdots,s\dot{q}^n(st))\right]$$

$$=\frac{d}{dt}s\sum_k g_{ik}(q^1(st),\cdots,q^n(st))\dot{q}^k(st)$$

$$=s^2\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}(q^1(\cdot),\cdots,q^n(\cdot),\dot{q}^1(\cdot),\cdots,\dot{q}^n(\cdot))\bigg|_{st}.$$

换句话说,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}(q^1(s\cdot),\cdots,\dot{q}^n(s\cdot)) - \frac{\partial L}{\partial q^i}(q^1(s\cdot),\cdots,s\dot{q}^n(s\cdot))\bigg|_t$$

$$=s^2\left[\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i}(q^1(\cdot),\cdots,q^n(\cdot)) - \frac{\partial L}{\partial q^i}(q^1(\cdot),\cdots,\dot{q}^n(\cdot))\right]\bigg|_{st} = 0.$$

因此曲线 $C_\xi(s\cdot)$ 还是此系统是轨线, 并且显然有

$$C'_\xi(s\cdot)|_0 = sC'_\xi(\cdot)|_0 = s\xi.$$

由微分方程的惟一性定理我们因而必定有

$$C_\xi(st) = C_{s\xi}(t). \tag{10.1}$$

这使我们定义一个映射, 它被称为 exp, 从 $T_p(M)$ 到 M,

$$\exp(\xi) = C_\xi(1). \tag{10.2}$$

注意, 映射 exp 在 $T_p(M)$ 原点的某个邻域内有定义和可微的. 事实上, 由 (10.1) 式,

$$\exp(\xi) = C_{\xi/\|\xi\|}(\|\xi\|),$$

其中 $\xi/\|\xi\|$ 现在位于 $T_p(M)$ 的单位球面上 [这是由 $T_p(M)$ 上的内积给出的欧氏度量下的单位球面]. 由于单位球面是紧的, 故存在某个 $\epsilon > 0$ 使得 $C_\eta(t)$ 对单位球面上所有的 η 和所有 $|t| < \epsilon$ 有定义. 因此 exp 对所有 $\|\eta\| < \epsilon$ 有定义.

映射 exp 是从向量空间 $T_p(M)$ 中原点的某邻域到此流形的可微映射. 让我们计算

$$\exp_{*0}: T_0[T_p(M)] \longrightarrow T_p(M).$$

设 $\xi \in T_p(M)$, 并考虑经过 $T_p(M)$ 中 0 的直线 l_ξ, 其定义为 $l_\xi(t) = t\xi$. 因为 $T_p(M)$ 为向量空间, 我们可以将 $T_0[T_p(M)]$ 经恒同分图表等同于 $T_p(M)$, 这时我们便可将 $l'_\xi(0)$ 与 ξ 等同. 但是

$$\exp[l_\xi(t)] = \exp(t\xi) = C_{t\xi}(1) = C_\xi(t), \tag{10.3}$$

并且这条曲线在 0 的切线正好是 ξ. 换言之,

$$\exp_{*0}[l'_\xi(0)] = \xi.$$

因此, 如果将 $T_0[T_p(M)]$ 等同于 $T_p(M)$, 则可断言

$$\exp_{*0}: T_0[T_p(M)] \rightarrow T_p(M)$$

由

$$\exp_{*0}(\xi) = \xi \tag{10.4}$$

给出. 特别, 映射 $\exp_{*0}$ 是非异的, 故由隐函数定理, 映射 exp 在原点的某个邻域中为微分同胚.

那么, 我们构造了从 $T_p(M)$ 的原点的某一邻域到 M 内的一个微分同胚映射 exp, 由 (10.3) 式它将通过原点的直线映到通过 p 点的轨线. 在这种无势能的情形中, 称此轨线为*测地线*, 其理由马上就会清楚. 如果以某个分图表 (U, α) 将 $T_p(M)$ 与 V 等同, 我们则可用 exp 引进一个新的分图表 (U', n_α), 其定义为

$$n_\alpha^{-1}(\xi_\alpha) = \exp(\xi).$$

分图表 n_α 具有性质 $n_\alpha(p) = 0$, 并且 n_α 将经过 p 的测地线带到经过原点的直线. 分图表 (U', n_α) 称为*测地正则分图表*, 而对应的坐标则被称作 M 上的*测地正则坐标*.

我们来考虑曲线 $C_\xi(\cdot) = \exp(\cdot\xi)$, 只要 $\|\xi\| < \varepsilon$, 它便在 $0 \leqslant t \leqslant 1$ 中有定义. 我们有

$$I[C_\xi(\cdot)] = \int_0^1 L(C'_\xi(t))dt = \frac{1}{2}\int_0^1 \|C'_\xi(t)\|^2 dt.$$

但是, 由能量守恒, 对于我们的系统的任一条轨线有 $H\circ\mathfrak{L}(C'(t))=$ 常数. 这时, 由于 $U=0, H\circ\mathfrak{L}(C'(t))=L(C'(t))=\frac{1}{2}\|C'_\xi(t)\|^2=$ 常数. 因为 $C'_\xi(0)=\xi$, 故 $\|C'_\xi(t)\|=\|\xi\|$, 因此

$$I[C_\xi(\cdot)]=\frac{1}{2}\int_0^1\|\xi\|^2dt=\frac{\|\xi\|^2}{2}. \tag{10.5}$$

现在设 $\{\beta_s\}$ 为 $T_p(M)$ 中的单参数旋转群. 于是

$$\varphi_s=\exp\circ\beta_s\circ\exp^{-1}$$

定义了在开集 $U=\exp\{v:\|v\|<\varepsilon\}\subset M$ 上的一个单参数群. 如果 $\|\xi\|<\varepsilon$, 由 (10.5) 式我们有

$$I[\varphi_s\circ C_\xi(\cdot)]=\frac{\|\beta_s\xi\|^2}{2}=\frac{\|\xi\|^2}{2},$$

所以由 (9.10) 式, 得到了

$$\frac{d}{ds}I[\varphi_s\circ C_\xi(\cdot)]=0=f_{Z_P}(\overline{C}(1))=\langle C'(1),Z_{C(1)}\rangle,$$

其中 Z 是 φ 的无穷小生成元. 但 $Z_{C(1)}=\exp_{*0}Y_\xi$, Y 为 $T_p(M)$ 上的向量场, 它是这个旋转单参数群的无穷小生成元。现在取任意的 β. 因此 Y_ξ 可以是切于 $T_p(M)$ 中半径为 $\|\xi\|$ 的球面的任意向量. 我们由此得到结论说, 对任意切于 $\exp S_{\|\xi\|}$ 的 η 有 $(\eta,C'_\xi(1))=0$, 其中的 $S_{\|\xi\|}$ 为 $T_p(M)$ 中半径为 $\|\xi\|$ 的球面 (见图 13.4). 换言之, 不仅通过原点的射线正交于 $T_p(M)$ 在欧几里得度量下的球面, 而且通过原点的一条射线在 exp 的像也正交于中心在 p 的在 M 的黎曼度量下一个球面的像.

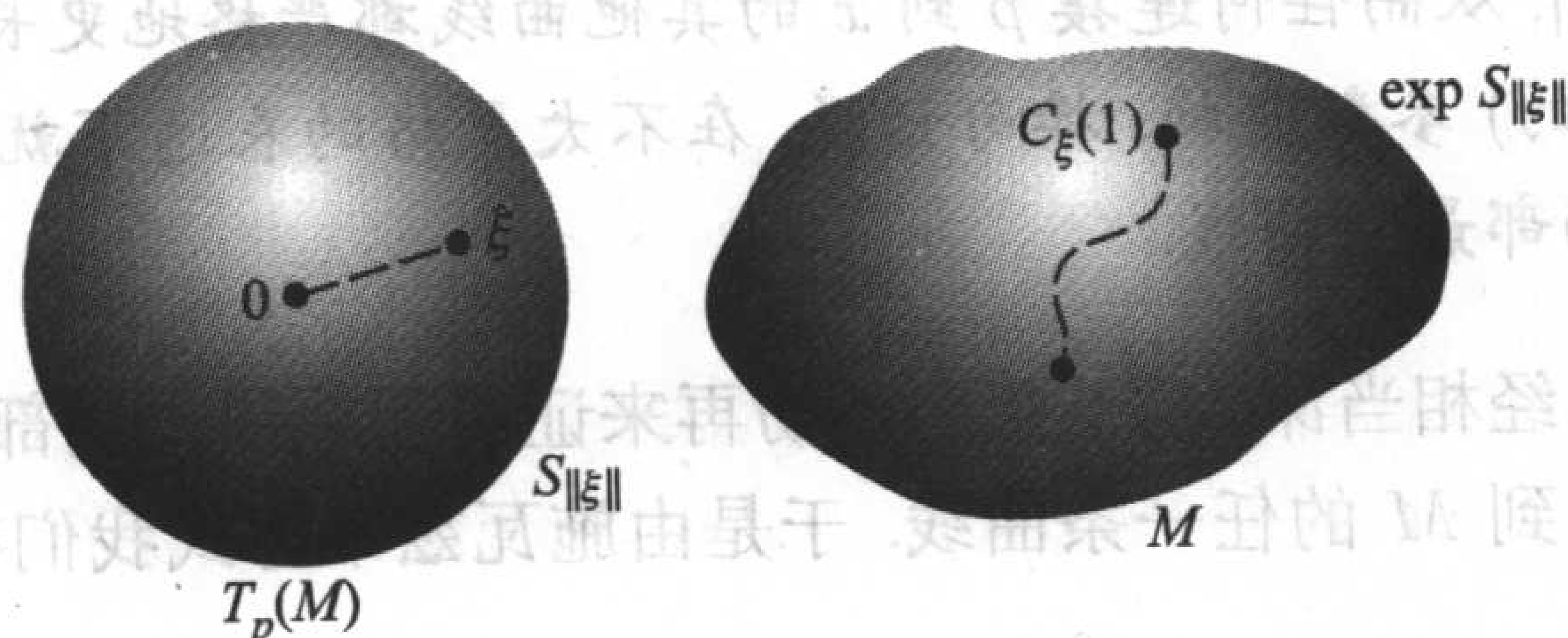


图 13.4

特别地, 我们可将 $T_p(M)$ 上的 "极坐标" 转移到 M 上, 从而得到 M 上的 "测地极坐标". 它的作用如下: 设 r 为 "径向坐标", 即 r 为在 U 中由

$$r(x)=\|\exp^{-1}x\|$$

定义的函数. 于是对任意 $x \in U, x \neq p$ 以及任意 $\zeta \in T_x(M)$ 我们有

$$\|\zeta\| \geqslant |\langle \zeta, dr \rangle|, \tag{10.6}$$

其中的等式只有当 ζ 切于过 p 的一条测地线时才成立. 事实上, 假设 $\zeta \in T_x(M)$, 其中的 $x = \exp\xi$ 而 $\xi \in T_p(M)$. 于是我们可以记 $\zeta = \zeta_1 + \zeta_2, \zeta_1$ 为 $C'_\xi(1)$ 的某个倍数, 而 ζ_2 切于 $\exp S_{\|\xi\|}$. 由上面的结果知 $(\zeta_1, \zeta_2) = 0$, 故

$$\|\zeta\|^2 = \|\zeta_1\|^2 + \|\zeta_2\|^2,$$

从而我们得到了 (10.6) 式, 其等式仅当 $\|\zeta_2\| = 0$ 时成立.

我们现在处在我们曾在第九章第 9 节处过的同样有利位置. 设 D 为联结从 p 到 x 的任一条曲线, 其中 $x = \exp\xi$. 设 t_1 是第一次使得 $D(t) \in \exp S_{\|\xi\|}$ 的 t. 于是

$$\begin{aligned} D\text{的长度} &= \int_0^1 \|D'(t)\|dt \geqslant \int_0^{t_1} \|D'(t)\|dt \geqslant \int_0^{t_1} |\langle D'(t), dr \rangle| \\ &\geqslant \int_0^{t_1} \langle D'(t), dr \rangle = r[D(t_1)] = \|\xi\|. \end{aligned}$$

另外, 等式仅当 $D'(t)$ 为一切于某条件 p 的固定测地线的切向量非负倍数时才成立. 于是 D 必为测地线 $C_\xi(\cdot)$. 简而言之, 我们已经证明了:

定理 10.1 设 $\varepsilon > 0$ 如此之小, 使得映射 exp 在 $B_\varepsilon = \{\xi \in T_p(M) : \|\xi\| < \varepsilon\}$ 上是个微分同胚. 设 $x = \exp\xi$ 是 $U = \exp B_\varepsilon$ 的一个点. 于是联结 p 到 x 的测地线 C_ξ 具有长度 $\|\xi\|$, 从而任何连接 p 到 x 的其他曲线都严格地更长, 除非 D 与 C_ξ 只差一个 (单调的) 参数变换. 换句话说, 在不太严格的表达下就是, 测地线是联结这两个点的局部最短曲线.

由于我们已经相当深入了, 我们不妨再来证明测地线也局部极小化了能量. 设 D 为从 $[0, 1]$ 到 M 的任一条曲线. 于是由施瓦兹不等式我们有

$$\left(\int_0^1 \|D'(t)\|dt \right)^2 \leqslant \int_0^1 \|D'(t)\|^2 dt \int_0^1 1dt = 2I(D),$$

等号仅当 $\|D'(t)\|$ 为常数时成立. 如果 $C_\xi(\cdot)$ 是连接 p 到 $x = \exp\xi$ 的测地线, 因此有

$$I[D] \geqslant \frac{1}{2} \left(\int_0^1 \|D'(t)\|dt \right)^2 \geqslant \frac{1}{2} \left(\int_0^1 \|C'(t)\|dt \right)^2 = \frac{1}{2}\|\xi\|^2.$$

现在在第二个不等式中的等号仅当 $D'(t)$ 与 $C'(t)$ 成比例时才成立, 而第一个不等式中仅当 $\|D'(t)\|$ 为常数时等号成立. 我们因此有结论 $\|D'(t)\| = \|\xi\|$, 即 $D'(t) = C'(t)$. 换句话说, 我们证明了下面的命题:

命题 10.1 *在定理 10.1 的假设下, $C_\xi(\cdot)$ 是在所有满足 $C(0) = p, C(1) = x$ 的曲线 $C : [0,1] \to M$ 中使 $I[C]$ 为严格的绝对极小的曲线.*

13.11 欧拉方程

在某种我们就要描述的条件下, 运动方程具有特别漂亮的形式. 我们要作的第一个特殊假设是存在一个从 $T(M)$ 到 $M \times V$ 的同构. 更准确地说, 我们假定有一个从 $T(M)$ 到 $M \times V$ 的微分同胚 ι, 使得 $\iota(\xi) = \prec m, v \succ$, 其中 $m = \pi(\xi)$, 并且对每个 $x \in M$, 映射 $\xi \mapsto v : T_x(M) \to V$ 是向量空间间的线性同构. 例如, 如果 M 是 V 的一个开子集, 那么等同分图表便定义了这样一个同构

$$\iota(\xi) = \prec \pi(\xi), \xi_{\mathrm{id}} \succ .$$

n 维的环面 $T^n = S^1 \times \cdots \times S^1$ 提供了一个不太平凡的例子. 那么我们可以引进"角变量" $\theta^1, \cdots, \theta^n$, 其中的 θ^i 为第 i 个圆上的角变量. 因此我们得到几个向量场 $\partial/\partial\theta^1, \cdots, \partial/\partial\theta^n$ 它们在 M 的每个点上线性无关的. 我们有了 $T_x(M)$ 的一组基, 因而有了 $T_x(M)$ 与 $\mathbb{R}^n = V$ 的一个同构, 它定义了所想要的映射 ι. 在下一节我们将遇到一个更复杂的例子. 应当指出, 只有极特殊类型的流形才具有这样一个从 $T(M)$ 到 $M \times V$ 的同构.

对本节剩下的部分, 我们把 $T(M)$ 与 $M \times V, T^*(M)$ 与 $M \times V^*$ 经由对应的 (相伴的) 同构等同.

把每个 $T_x(M)$ 与 V 等同的这个规则可以看为 M 上的一个 V - 值微分形式. 以 ω 表示此形式. 换而言之, 等同 ι 对 $\xi \in T_x(M)$ 由 $\iota(\xi) = \langle \xi, \omega_x \rangle$ 给出. 因此我们可以研究 V - 值外 2 - 形式 $d\omega$. 对每一对切向量 $\xi, \eta \in T_x(M)$, 我们得到 $\langle \eta, \xi \rfloor d\omega \rangle$ 为 V 中的一个元素. 现在可以将 ξ 和 η 等同于 V 中的向量. 因此我们得到 V 上的一个 V - 值反对称双线性形式; 如果称其为 $\mathfrak{a}_x$, 则有

$$\mathfrak{a}_x(v, w) = \langle \eta, \xi \rfloor d\omega \rangle,$$

其中 $\langle \xi, \omega_x \rangle = v, \langle \eta, \omega_x \rangle = w, \xi, \eta \in T_x(M)$. 注意, 一般地, 这个双线性形式 $\mathfrak{a}_x$ 依赖于 x. 我们关于等同映射 ι 的第二个基本假定是说 $\mathfrak{a}_x$ 与 x 无关. 就是说, 假设

存在一个 V - 值双线性形式 $\mathfrak{a}$, 使得

$$\langle Y, X \rfloor d\omega \rangle = \mathfrak{a}(\langle X, \omega \rangle, \langle Y, \omega \rangle) \tag{11.1}$$

对 M 上所有向量场 X 和 Y 成立.

在上面所给出的例子 (即环面上的开子集) 中 $d\omega = 0$, 故 (11.1) 式显而易见成立. 下一节我们将遇到 $d\omega \neq 0$ 的情形.

为了更好一些理解 (11.1) 式, 我们引进下面的记号: 对任意 $v \in V$, 设 $\hat{v}$ 为 M 上由 $\langle \hat{v}, \omega \rangle_x = v$ 给出的向量场, 其中 $x \in M$ 为所有的点, 于是对任意 $v, w \in V$ 我们有

$$\langle \hat{w}, \hat{v} \rfloor d\omega \rangle = D_{\hat{v}} \langle \hat{w}, \omega \rangle - D_{\hat{w}} \langle \hat{v}, \omega \rangle - \langle [\hat{v}, \hat{w}], \omega \rangle.$$

现在 $\langle \hat{w}, \omega \rangle = w, \langle \hat{v}, \omega \rangle = v$ 为常值, 故上式右端的头两项为零. 因而可重写 (11.1) 式为

$$\mathfrak{a}\langle v, w \rangle = -\langle [\hat{v}, \hat{w}], \omega \rangle. \tag{11.1'}$$

对于一个力学系统的动能我们需要 M 上的一个黎曼度量. M 上的一个黎曼度量给出了在每个 $T_x(M)$ 上的一个内积. 这意味着对每个 $x \in M$ 给出了 V 上的一个内积 $(\ ,\ )_x$. 我们的第三个特殊的假设是 $(\ ,\ )_x$ 不依赖于 x. 因而我们有了 V 上的一个内积, 它通过 $T(M)$ 与 $M \times V$ 的等同给出了 M 上的黎曼度量.

我们希望把 $T^*(M)$ 上的向量场描述为 X_{-dH}, 其中的 $H = K + U, K$ 为黎曼度量的动能, U 为某个势能. 因为 $T^*(M) = M \times V^*, T^*(M)$ 上的一个向量场 X 可以被惟一地写成 $X = X_1 + X_2$, 其中 X_1 切于 M 而 X_2 切于 V^*. 另外, 我们可以将 X_1 看成一个 V - 值函数而 X_2 为一个 V^* - 值函数 (将向量空间 V^* 的切空间等同于 V^*). 于是

$$\langle X, \theta \rangle_{\prec x, l \succ} = \langle X_1, l \rangle$$

对任意 $\prec x, l \succ \in M \times V^*$ 成立. 我们其实应该真正地把它写成

$$\langle \cdot, \theta \rangle = \langle \langle \cdot, \pi^* \omega \rangle, p \rangle,$$

其中 ω 为前面所定义的形式, 而 V^* - 值函数 $p: M \times V^* \to V^*$ 是到第二个因子的投影, $p(m, l) = l$. 于是

$$\langle \cdot, X \rfloor d\theta \rangle = \langle \langle \cdot, \pi^* \omega \rangle, \langle X, dp \rangle \rangle - \langle \langle X, \pi^* \omega \rangle, \langle \cdot, dp \rangle \rangle - \langle \langle \cdot, X \rfloor \pi^* d\omega \rangle, p \rangle.$$

作替换 $Y = Y_1 + Y_2$ 并利用 (11.1) 式, 我们得到

$$\langle Y, X \rfloor d\theta \rangle = \langle Y_1, X_2 \rangle - \langle X_1, Y_2 \rangle - \langle \mathfrak{a}(X_1, Y_1), l \rangle.$$

现在 $H = K + U$, 其中 $K(l) = \frac{1}{2}(l, l), U(x, l) = \overline{U}(x)$. 因此

$$\langle Y, dH \rangle = (Y_2, l) + Y_1 \overline{U},$$

从而方程

$$\langle Y, X \rfloor d\theta \rangle = -\langle Y, dH \rangle$$

成为

$$\langle Y_1, X_2 \rangle - \langle X_1, Y_2 \rangle - \langle \mathfrak{a}(X_1, Y_1), l \rangle = -(Y_2, l) - \langle Y_1, d\overline{U} \rangle, \tag{11.2}$$

它必然对 Y 的所有选取都成立.

令 $Y_1 = 0$, 我们得到

$$\langle X_1, \cdot \rangle = (\cdot, l). \tag{11.3}$$

事实上, 出现在最后面这个等式中的内积是在 V^* 上的. 把它转移为 V 上的等式, 我们得到

$$(X_1, \cdot) = \langle \cdot, l \rangle. \tag{11.4}$$

令 $t \mapsto \prec C(t), v(t) \succ$ 是把我们的系统用黎曼度量转移到 $M \times V = T(M)$ 后的解曲线. 那么最后面的那个等式是说 $C'(t) = v(t)$.

如果在 (11.2) 式中令 $Y_2 = 0$, 并运用 (11.4) 式便得到

$$\langle \cdot, X_2 \rangle - (\mathfrak{a}(X_1, \cdot), X_1) = -\langle \cdot, d\overline{U} \rangle$$

现在对 (C, v) 的任何解我们都有

$$\langle \cdot, X_2 \rangle = \left(\cdot, \frac{dv}{dt} \right),$$

所以最后面的这两个方程可以重写为

$$\begin{aligned} C'(t) &= v(t), \\ \left(\cdot, \frac{dv}{dt} \right) &= (\mathfrak{a}(v, \cdot), v) - \langle v, d\overline{U} \rangle, \end{aligned} \tag{11.5}$$

它被称作欧拉方程.

13.12 刚体运动

我们要把前一节的结果用于刚体运动的研究. 为简明起见, 我们只局限于具有一个固定点的刚体运动的研讨. 刚体整个移动的更一般的情形可由相似的方法掌握.(经常是首先考虑重心的运动, 而更一般的情形可分成两部分: 重心的运动和物体相对于重心的运动. 于是它化成了我们正要研究的问题.)

为了说清我们的方法的普遍性, 我们考虑在 $\mathbb{E}^n$ 中一个刚体的运动方程. 只在本节末尾才用到 $n=3$. 我们在物体中 "作通过此固定点的" 某个正的法正交系, 并固定它. 换句话说, 固定了物体的一个初始位置 $x_0 = \prec b^1, \cdots, b^n \succ$. 物体的任意其他的位置 x_1 可以由 x_0 经过旋转得到: $x_1 = R_1 x_0$. 设 $R(t) = e^{At}$ 为一个旋转的单参数群. 于是

$$R_1 R(t) x_0 = R_1 R(t) R_1^{-1} x_1$$

是此刚体的可能的位置曲线. 这条曲线在 x_1 的切线被记为 $\hat{A}_{x_1}$. 因而 $\hat{A}_{x_1} \in T_{x_1}(M)$. 如果 $x_2 = R_2 x_0 = R_2 R_1^{-1} R_1 x_0 = R_2 R_1^{-1} x_1$, 因此 $\hat{A}_{x_2}$ 切于曲线 $R_2 R(t) x_0 = R_2 R_1^{-1} R_1 R(t) x_0$, 从而 $\hat{A}_{x_2} = (R_2 R_1^{-1})_* \hat{A}_{x_1}$. 由定义显见 $\hat{A}_{x_1} = 0$ 当且仅当 $A = 0$.

我们可以把映射 $A \mapsto \hat{A}_x$ 看作从斜伴随线性变换的空间到 $T_{x_1}(M)$ 的映射. 设 V 表示斜伴随线性变换的空间. 于是由于

$$\dim V = \dim M = \frac{n(n-1)}{2},$$

并且映射 $A \mapsto \hat{A}_{x_1}$ 为单射, 我们得出它是个同构的结论. 我们因此有了一个平凡化映射 $T(M) \mapsto M \times V$. 因此, 像在第 11 节那样, 我们得到了一个 V - 值线性微分形式 ω.

让我们再一次描述 $\omega_x : T_x(M) \to V$ 的意义. 如果 $\xi \in T_x(M)$ 代表了此刚体的无穷小运动, 那么, 因为此物体是刚性的, 我们可以将 ξ 看作此物体相对于物体外的一个观察者 (固定于空间中) 的一个无穷小旋转. 因而 $\xi = B$, 其中譬如是某个 $B \in V$. 于是 $\langle \xi, \omega \rangle$ 是在附加在此物体上的基下表达出的对应的无穷小旋转; 就是说, $\langle \xi, \omega \rangle = R_1^{-1} B R_1$, 其中 $x_1 = R x_0$.

用 $\hat{A}$ 为对应于 $A \in V$ 的一个向量场 $x \mapsto \hat{A}_x$. 设 φ_t 为由 $\hat{B}$ 生成的单参数群, 其中 $B \in V$. 注意, 在任意 $x_2 = R_2 x_0$, 我们有 $\varphi_t x_2 = R_2 e^{Bt} x_0$. 于是在任意 $x_1 = R_1 x_0$ 我们有

$$\varphi_t R_1 e^{At} R_1^{-1} x_1 = \varphi_t R_1 e^{At} x_0 = R_1 e^{At} e^{Bt} x_0 = R_1 e^{Bt} (e^{-Bt} e^{At} e^{Bt}) x_0,$$

所以

$$\varphi_{t*}\hat{A}_{x_1} = (\widehat{e^{-Bt}Ae^{Bt}})_{\varphi_t x_1}.$$

如果在 $t=0$ 处对此等式做关于 t 的微分, 我们得到

$$[\hat{A}, \hat{B}] = (AB - BA) = -[A, B],$$

所以按照 (11.1) 式, 我们有

$$\mathfrak{a}(A, B) = [A, B] = BA - AB. \tag{12.1}$$

我们现在要指出在此物体上的一个质量分布如何决定了 M 上的一个黎曼度量. 设 p 为此物体上一个质量为 m 的质点, 假设质点 p 有坐标 $\prec p^1, \cdots, p^n \succ$, 这是对于物体上所作的那些轴而言的. 假设物体处在 $x = \prec b_1, \cdots, b_n \succ$ 位置. 于是质点 p 位于点 $p^1 b_1 + \cdots + p^n b_n \in \mathbb{E}^n$. 如果 $R(t) = e^{At}$ 为一个旋转单参数群, 则当物体在 $R_1 R(t) x$ 时, 质点 p 则位于

$$R_1[p^1 R(t) b_1 + \cdots + p^n R(t) b_n] = R_1 R(t) p.$$

因而当物体进行由 $\hat{A}$ 生成的运动时, 质点 p 的速度为 $R_1 Ap$, 而质点 p 的动能是 $\frac{1}{2}m\|R_1 Ap\|^2 = \frac{1}{2}m\|Ap\|^2$, 这是因为 R_1 是个正交线性变换. 定义 $\hat{A}_x$ 的动能为此物体所有质点的总动能. 因而

$$\frac{1}{2}\|\hat{A}_x\|^2 = \frac{1}{2}\int_{\text{物体}} m\|Ap\|^2. \tag{12.2}$$

注意, $\|\hat{A}_x\|$ 只依赖于 A, 从而假定 (12.2) 式的确定义了 V 上的一个范数, 则我们在上一节中的第三个要求得到满足.(注意: 质量的分布可以对所有的 $p \in \{p : m(p) > 0\}$ 有 $Ap = 0$, 但这不表示 $A = 0$. 例如, 假设所有的质量都集中于一条直线 l. 如果 A 代表绕 l 的无穷小旋转, 故对 $p \in l$ 有 $Ap = 0$, 于是 $\|A\| = 0$. 但是, 如果对那些具有 $m > 0$ 的 p 的集合张成了 $\mathbb{E}^n$, 显然 (12.2) 式定义了一个黎曼度量. 事实上, 对所有属于一个张成集合的 p

$$\|A\| = 0 \Rightarrow Ap = 0,$$

因此 $A = 0$).

让我们更仔细一点考查 (12.2) 式给出的内积. 如果 f 表示 $\mathbb{E}^n \otimes \mathbb{E}^n$ 上对应于 $\mathbb{E}^n$ 上内积的线性函数 (它是个双线性形式); 换句话说,

$$f(a_1 \otimes b_1 + \cdots + a_k \otimes b_k) = (a_1, b_1) + \cdots + (a_k, b_k).$$

设 s 为 $V \otimes V$ 中的任意元素. 于是 s 定义了 Hom ($\mathbb{E}^n$) 上的一个双线性形式:

$$s(A, B) = f((A \otimes B)s).$$

注意, 如果张量 s 是对称的, 那么对应的双线性形式也是. 在当前的情形中, 内积 (12.2) 式来自张量 $I \in V \otimes V$, 其中 $I = \int_{\text{物体}} mp \otimes p$ 被称作这个**物体的惯性张量**.[1]因而 (12.2) 式可以写为

$$\|\hat{A}\|^2 = I(A, A).$$

这时的欧拉方程成为

$$C'(t) = A(t)\text{和}\left(\cdot, \frac{dA}{dt}\right) = I([\cdot, A], A) - \langle \hat{A}_{C(t)}, d\overline{U} \rangle. \tag{12.3}$$

现在张量 I 是对称的. 我们因此可以找到一个 $\mathbb{E}^n$ 的法正交基 $e^1, \cdots, e^n$, 使 I 化为对角形, 故

$$I = I_1 e^1 \otimes e^1 + I_2 e^2 \otimes e^2 + \cdots + I_n e^n \otimes e^n.$$

设 $E_{ij}(i < j)$ 为下面所定义的反称矩阵:

$$E_{ij}e^i = e^j, E_{ij}e^j = -e^i, E_{ij}e^l = 0(l \neq i, j).$$

于是

$$I(E_{ij}, E_{kl}) = 0, \text{当} i, j \neq k, l$$

以及

$$I(E_{ij}, E_{ij}) = I_i + I_j.$$

现在让我们看看方程 (12.3) 对 $n = 3$ 的情形说了些什么, 这时我们有

$$[E_{12}, -E_{13}] = E_{23}, [E_{12}, E_{23}] = -E_{13}, [E_{13}, E_{23}] = E_{12}.$$

假设 $A(t) = a_1(t)E_{23} - a_2(t)E_{13} + a_3(t)E_{12}$, 并设

$$B = b_1E_{23} - b_2E_{13} + b_3E_{12} = \text{常数}.$$

[1]这与物理教科书中通常叫作的惯性张量略有不同. 利用下面引进的系数 I_i, 通常叫做 I_{ij} 的在我们这里是 $I_i + I_j$.

于是将 B 代入 (12.3) 式并比较 b_1, b_2, b_3 的系数, 我们得到了

$$
\begin{aligned}
(I_2+I_3)\frac{da_1}{dt} &= (I_3-I_2)a_2a_3 + a_1\langle E_{12}, d\overline{U}\rangle, \\
(I_1+I_3)\frac{da_2}{dt} &= (I_1-I_3)a_1a_3 + a_2\langle -E_{12}, d\overline{U}\rangle, \\
(I_1+I_2)\frac{da_3}{dt} &= (I_2-I_1)a_1a_2 + a_3\langle E_{12}, d\overline{U}\rangle.
\end{aligned}
\tag{12.4}
$$

在没有势能的情形即 $U=0$ 下所产生的这些方程是一个简单的情形. 首先, 假定物体有一个球面对称的质量分布. 于是 $I_1=I_2=I_3$, 并且方程 (12.4) 成为

$$\frac{da_1}{dt} = \frac{da_2}{dt} = \frac{da_3}{dt} = 0$$

换句话说, $A=$ 常数. 该物体的运动由绕固定于物体上的某个轴的均匀旋转构成. 当然这一方法在空间上作为一个观察者物体也受到绕一固定轴均匀旋转.

其次, 我们考虑自由运动的一个轴对称刚体. 即 $I_1=I_2$, 且 $U=0$. 于是运动方程成为

$$
\begin{aligned}
\frac{da_1}{dt} &= Ka_2a_3, \\
\frac{da_2}{dt} &= -Ka_1a_3 \\
\frac{da_3}{dt} &= 0,
\end{aligned}
$$

其中

$$K = \frac{I_3-I_2}{I_3+I_2}.$$

这些方程的解可以立刻写出来: $a_3 = s =$ 常数, 而

$$
\begin{aligned}
a_1 &= c_1\cos Kst + c_2\sin Kst, \\
a_2 &= -c_1\sin Kst + c_2\cos Kst.
\end{aligned}
$$

这个运动以正则进动而知名.(不应将这个运动与地轴的天文进动相混淆, 这是由于太阳和月亮的万有引力作用产生的.)

如果 I_1, I_2, I_3 中没有两个是相等的, 则这个运动方程仍可用积分解出, 只是表达式十分复杂. 建议读者参阅任何一本关于力学的标准读本以了解其细节.

到目前为止, 我们一直在考虑无势能的运动. 如果 $\overline{U}\neq 0$, 则方程 (12.4) 通常不那么容易就解的. 我们来处理一个对称陀螺的情形, 即 $I_1=I_2$, 而 U 由重

力势能给出. 为了解这个问题, 使用天文学中的欧拉角作为 M 上的局部坐标是方便的. 为避免混淆, 我们重新作出定义: 设 $\delta_x, \delta_y, \delta_z$ 为 $\mathbb{E}^3$ 的基向量, 它对应于以物体上的固定点为原点的直角坐标 x, y, z. 我们可假定此刚体的质心与这个固定点不同, 否则就不会有重力作用于此物体了. 容易看出, 质点 c 位于对称轴上. 我们假定在此物体上所作的向量使 b_3 从固定点指向质心, 即从 0 到 c. 于是 $0 < \theta < \pi$ 被定义为 b_3 和 δ_z 之间的夹角:

$$\cos\theta = (b_3, \delta_z).$$

结点线是由 b_1 和 b_2 张成的平面与 xy 平面的交线.(因此为了使它有定义, 我们心须限制于 $b \neq \pm\delta_z$ 这个开集.) 我们定义沿结点线的单位向量 n 为那个使 δ_z, b_3, n 构成一个正的 (即右手系)$\mathbb{E}^3$ 的基.

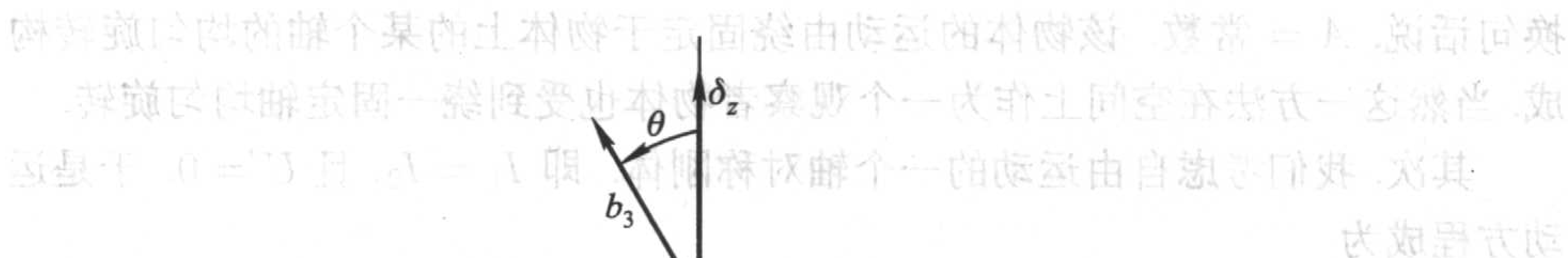


图 13.5

角 $0 < \psi < 2\pi$ 定义为 n 与 δ_x 构成的角, 而 $0 < \varphi < 2\pi$ 为 n 与 b_1 间的角. 我们现在希望找到联系 $T_x(M)$ 由 $\partial/\partial\theta, \partial/\partial\varphi, \partial/\partial\psi$ 给出的基与早先引进的正交基 $E_{12}, -E_{13}, E_{23}$ 的变换. 假设 x 有坐标 $\prec \theta, \psi, \varphi \succ$(见图 13.5). 现在 $(\partial/\partial\theta)_x$ 代表了绕结点线的无穷小旋转, 而 $n = (\cos\varphi)b_1 - (\sin\varphi)b_2$, 所以有

$$\left(\frac{\partial}{\partial\theta}\right)_x = (\cos\varphi)E_{23} - (\sin\varphi)(-E_{13}).$$

向量 $(\partial/\partial\varphi)_x$ 代表了绕 b_3 的无穷小旋转, 所以

$$\left(\frac{\partial}{\partial\varphi}\right)_x = E_{12}.$$

最后, $(\partial/\partial\psi)_x$ 代表了绕 δ_z 的无穷小旋转. 现在

$$(\delta_z, b_3) = \cos\theta.$$

另外, 由于 $(\delta_z, n) = 0$, 则 δ_z 到由 δ_1 和 δ_2 张成的平面上的投影必定仍旧正交于 n, 这是因为 n 也在此平面中. 因此容易验证这个投影由 $(\sin\varphi)b_1+(\cos\theta)b_2$, 因而

$$\delta_z = \sin\theta[(\sin\varphi)b_1 + (\cos\varphi)b_2] + (\cos\theta)b_3,$$

于是

$$\left(\frac{\partial}{\partial\psi}\right)_x = (\sin\theta\sin\varphi)E_{23} + (\sin\theta\cos\varphi)(-E_{12}) + (\cos\theta)E_{12}.$$

如果 $\xi\in T_x(M)$ 由

$$\xi = \dot{\theta}\left(\frac{\partial}{\partial\theta}\right)_x + \dot{\psi}\left(\frac{\partial}{\partial\psi}\right)_x + \dot{\varphi}\left(\frac{\partial}{\partial\varphi}\right)_x$$

给出, 我们则有

$$\begin{aligned}2K(\xi) = \|\xi\|^2 = &\dot{\theta}^2(\cos\theta)^2(I_2+I_3) + (\sin\varphi)^2(I_1+I_3) + \dot{\varphi}^2(I_1+I_2)\\ &+2\dot{\varphi}\dot{\psi}(I_1+I_2)\cos\theta + \dot{\psi}^2\{(\sin^2\theta)[(I_2+I_3)(\sin)^2\varphi\\ &+(I_1+I_3)\cos^2\varphi] + (I_1+I_2)\cos^2\theta\}\end{aligned}$$

由假设 $I_1 = I_2$, 令

$$M_1 = I_1 + I_3 = I_2 + I_3, M_2 = I_1 + I_2,$$

则上面的表达式简化为

$$K(\theta,\psi,\varphi;\dot{\theta},\dot{\psi},\dot{\varphi}) = \frac{1}{2}M_1(\dot{\theta}^2 + \dot{\psi}^2\sin^2\theta) + \frac{1}{2}M_2(\dot{\varphi} + \dot{\psi}\cos\theta)^2. \tag{12.5}$$

我们来得出势能的表达式. 如果我们假定 (我们将要如此) 有了一个均匀的垂直的引力场, 它与重心的高成正比. 因而

$$\overline{U}(\theta,\psi,\varphi) = k\cos\theta, \tag{12.6}$$

其中 $k = mg\|c\|$, m 为物体的总质量, g 是重力场的力, $\|c\|$ 是从固定点到重心的距离. 于是拉格朗日方程由

$$L(\theta,\psi,\varphi;\dot{\theta},\dot{\psi},\dot{\varphi}) = \frac{1}{2}M_1(\dot{\theta}^2 + \dot{\psi}^2\sin^2\theta) + \frac{1}{2}M_2(\dot{\varphi} + \dot{\psi}\cos\theta)^2 - k\cos\theta. \tag{12.7}$$

注意, $\partial/\partial\varphi$ 和 $\partial/\partial\psi$ 都是黎曼度量下的等度映射, 使得 $\overline{U}$ 不变, 因此对应的动量是此运动的常数. 换句话说, 对此系统的任意运动我们有

$$P_{\partial/\partial\psi} = C_1, P_{\partial/\partial\varphi} = C_2,$$

其中 C_1, C_2 为常量. 但是

$$P_{\partial/\partial\psi} = \frac{\partial L}{\partial \dot{\psi}} = M_1\dot{\psi}\sin^2\theta + M_2\cos\theta(\dot{\varphi} + \dot{\psi}\cos\theta) = C_1,$$
$$P_{\partial/\partial\varphi} = M_2(\dot{\varphi} + \dot{\psi}\cos\theta) = C_2.$$

解出这些方程得到

$$\dot{\varphi} + \dot{\psi}\cos\theta = \frac{C_2}{M_2}, M_1\dot{\psi}\sin^2\theta = C_1 - C_2\cos\theta. \tag{12.8}$$

把 (12.8) 式代入能量表达式

$$E = K + \overline{U} = \frac{1}{2}M_1(\dot{\theta}^2 + \dot{\psi}^2\sin^2\theta) + \frac{1}{2}M_2(\dot{\varphi} + \dot{\psi}\cos\theta) + k\cos\theta,$$

则对给定的 C_1 和 C_2 值得到了表达式

$$\frac{C_2^2}{2M_2} + \frac{1}{2}M_1\dot{\theta}^2 + \frac{(C_1 - C_2\cos\theta)^2}{2M_1\sin^2\theta} + k\cos\theta = 常数. \tag{12.9}$$

因而如同我们对中心力问题的处理那样, 对 $P_{\partial/\partial\varphi}$ 和 $P_{\partial/\partial\psi}$ 的固定值, θ 的运动能量被 (12.9) 式的一个一维力学系统决定. 由这个力学系统中解出 θ 后, 我们可以从 (12.8) 式通过积分得到 $\psi(\cdot)$ 和 $\varphi(\cdot)$. 为了得到对解 θ 的行为的定性的信息, 我们则可把第 5 节的方法用到我们的一维力学系统. 注意, 如果 $C_1 \neq C_2$ (物体的 "快速旋转" 便会是这种情形) 则当 $\theta \to 0$ 和当 $\theta \to \pi$ 时动能趋向于无穷. 因此我们得到 θ 在两个值 $0 < \theta_1 \leqslant \theta \leqslant \theta_2 < \pi$ 中摆动. 换句话说, 该物体的对称轴进行一个周期的往复运动 (称其为*章动*).当 θ 振动时 ψ 满足 (12.8) 式. 让我们画出 b_3 在单位球面上勾画出的曲线. 如果 $C_1 > C_2\cos\theta > 0$ 且 $\theta_1 \leqslant \theta \leqslant \theta_2$, 则虽然 ψ 的大小值在摆动, 但有 $\psi > 0$. 因而此运动如图 13.6 所展示的那样. 另一种可能性是

$$C_1 - C_2\cos\theta_1 < 0 \quad 而 \quad C_1 - \cos\theta_2 > 0,$$

从而在 $\theta = \theta_1$ 附近 ψ 为负, 而在最大的 $\theta = \theta_2$ 附近为正. 这时在一个周期上 ψ 的平均值仍是正的, 故此运动如同在图 13.7 中展示的那样.

极限情形是 $C_1 - C_2\cos\theta_1 = 0$, 这时 b_3 的运动如同图 13.8 那样.(如果一个旋转陀螺的轴固定在某个位置 θ, ψ 并且随后倒下, 就是这种情形.)

在所有这些情形中物体轴绕 z 轴的运动都被称作进动. 应该记住该陀螺在我们的讨论中一直以常值角动量 C_2 在绕它的轴旋转.

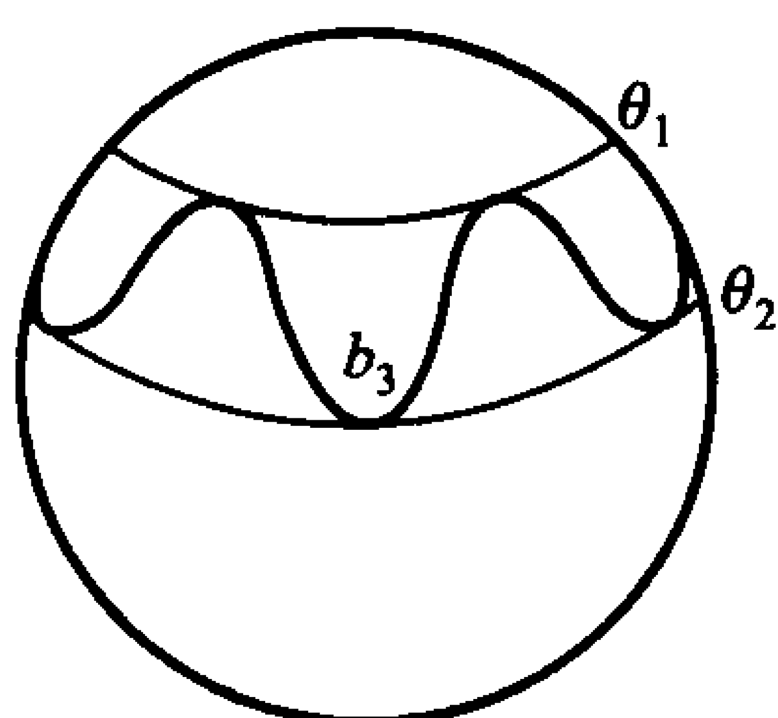


图 13.6

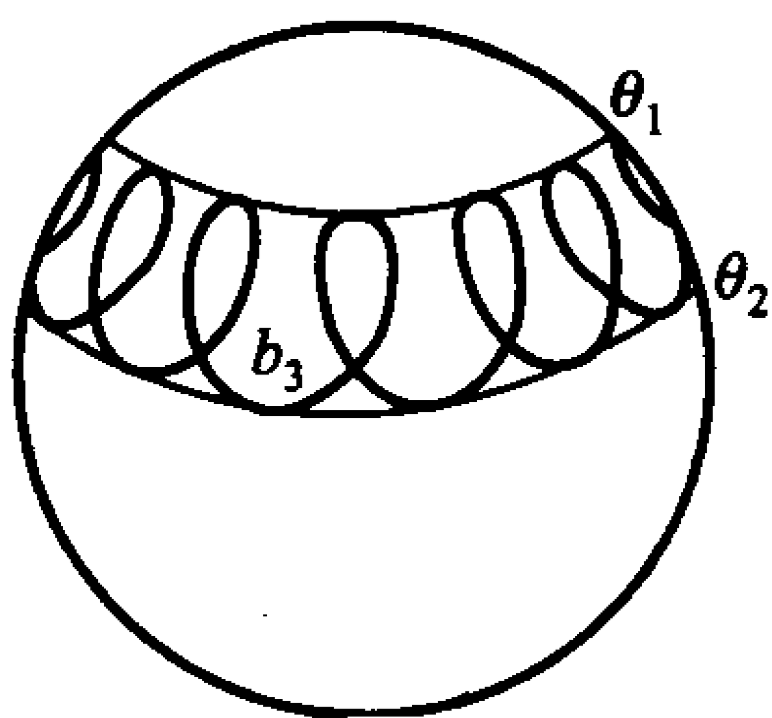


图 13.7

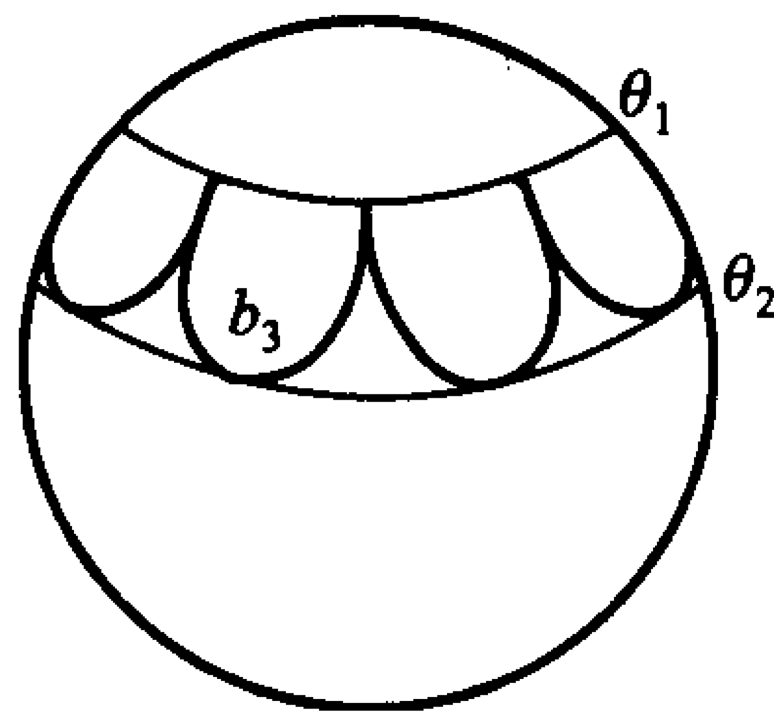


图 13.8

13.13 小振动

假设我们有一个流形 M 上的力学系统, 其能量为 $H = K + U$. 假设 "力场" $d\overline{U}$ 在某点 $x_0 \in M$ 为零. 于是常值曲线 $C(t) \equiv x_0$ 是此系统的一条轨线. 事实上, 选取一个分图表 (W, α), 其坐标为 $x^1, \cdots, x^n$, 使得

$$\alpha(x_0) = \prec 0, \cdots, 0 \succ.$$

于是在 $\pi^{-1}(W)$ 上的对应坐标 $\prec q^1, \cdots, q^n \succ$ 中, 我们在点 $\prec 0, \cdots, 0, \cdots, 0 \succ$ 有

$$\frac{\partial H}{\partial q^i} = \frac{\partial \overline{U}}{\partial x^i} = 0 \quad 和 \quad \frac{\partial H}{\partial p^i} = \frac{\partial K}{\partial p^i} = 0.$$

因此自然期望对于 q 和 p 的小的初值和小的时间区间, 该系统的解应该被下面的线性系统很好地逼近:

把势能

$$\overline{U}(x^1, \cdots, x^n) = \sum a_{ij} x^i x^j + U_3$$

[其中 $U_3 = \mathfrak{O}(\|x\|^3)$; 即 U_3 在 $x = 0$ 以 3 阶取 0]换为二次的势能

$$\overline{U}_\alpha(x) = \frac{1}{2} \sum a_{ij} x^i x^j,$$

并把对应于已知黎曼度量的动能

$$K(q, \dot{q}) = \frac{1}{2} \sum g_{ij}(q) \dot{q}^i \dot{q}^j,$$

换为对应于在 x_0 的欧几里得度量的动能

$$K_2(q, \dot{q}) = \frac{1}{2} \sum g_{ij}(0, \cdots, 0) \dot{q}^i \dot{q}^j.$$

因此我们得到了一个力学系统 H_2, 其对应的方程 (5.1) 实际上是线性的. [读者应该将此作为习题去验证这些方程完全是向量场 X_{-dH} 沿曲线 $\overline{C}(t) = \prec 0,\cdots,0 \succ \in T^*(M)$ 的变分 (在第 2 节中介绍过的)].

当然, 当时间增大时, q^i 和 p^i 的值会变得相当大, 从而线性逼近无效, 但是, 在某些情形下, 我们对所有时间可以确保 q^i 和 p^i 一直是小的. 事实上, 假设二次形式 $\sum a_{ij}x^i x^j$ 是正定的. 于是 $\overline{U}$ 在 x_0 有一个严格的极小值, 譬如 $\overline{U}(x_0) = 0$. 特别, 如果我们在 x_0 以动能 $K = E$ 开始, 其中 E 为充分小, 则由能量 $K + \overline{U} \equiv E$ 守恒知 x 被限制在由 $\overline{U}(x) \leqslant E$ 定义的一个 x_0 的邻域中, 而动量被限制在条件 $K \leqslant E$ 下 (见图 13.9). 因而 q^i 和 p^i 保持为小的值. 这并不意味着具哈密顿 H 的原来力学系统的解将与该线性系统的一个固定的解曲线一直都接近. 它的意思是说对任何短小的时间区间 (即任何未来时间附近的一个短小区间), 该轨线将靠近此线性系统的某条轨线.

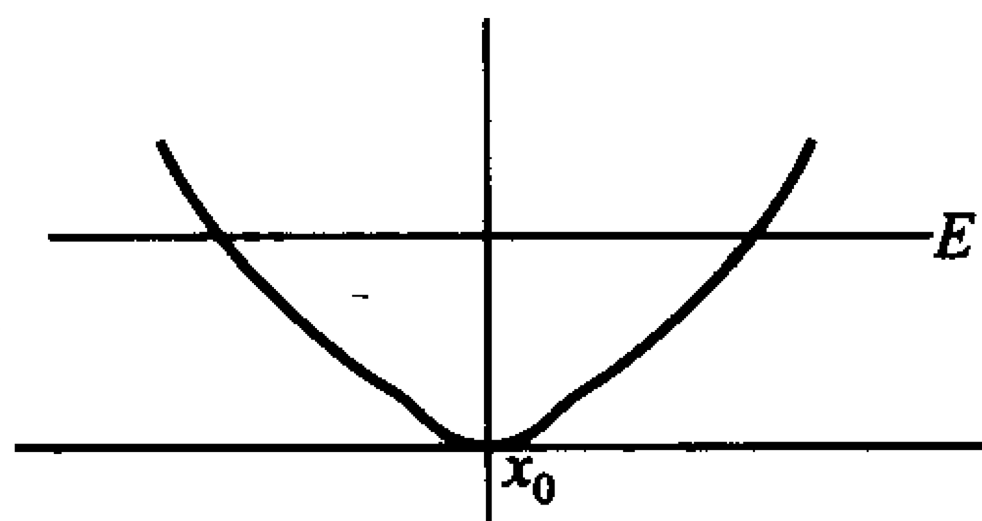


图 13.9

因而我们对下面类型的力学系统感兴趣: 在某个向量空间中的构形空间 M. 在该流形上的黎曼度量由欧几里得度量给出. 势能是此向量空间上的一个正定二次形式. 假设关于此欧氏度量我们选取了直角坐标 $x^1,\cdots,x^n$. 因而

$$\|(\dot{q}^1,\cdots,\dot{q}^n)\|^2 = (\dot{q}^1)^2 + \cdots + (\dot{q}^n)^2,$$

而映射 $\mathfrak{L}$ 由

$$p^i = \dot{q}^i$$

给出, 因此

$$L(q,\dot{q}) = K(\dot{q}) - \overline{U}(q) = \frac{1}{2}\left(\sum(\dot{q}^i)^2 - \sum a_{ij}q^i q^j\right)$$

而拉格朗日方程为

$$\frac{dq^i}{dt} = \dot{q}^i, \frac{d\dot{q}^i}{dt} = -\sum a_{ij}q^j \tag{13.1}$$

(当然, 这正是现在 $\mathfrak{a} \equiv 0$ 情形的欧拉方程 (11.15).)

方程 (13.1) 可以写得更具启发性: 设 A 为矩阵是 (a_{ij}) 的线性变换. 以不变性的语言来说, 就是 A 为惟一的自伴线性变换使得

$$\overline{U}(x) = (Ax, x). \tag{13.2}$$

于是该系统的轨线 $v(\cdot)$ 是二阶微分方程

$$\frac{d^2v}{dt^2} = -Av \tag{13.3}$$

的解.

要求 (13.1) 和 (13.3) 式的实际解, 我们可利用第五章的定理 3.1. 根据那个定理, 如果 M 是有限维的, 我们则可找到 A 的特征值的一组法正交基 $e^1, \cdots, e^n$. 换句话说, 如果 z^i 是对应于此基的直角坐标, 则

$$\overline{U}(z^1, \cdots, z^n) = \lambda_1(z^1)^2 + \cdots + \lambda_n(z^n)^2,$$

从而方程 (13.1) 和 (13.3) 成为

$$\frac{d^2z^i}{dt^2} = -\lambda_i z^i.$$

因此 (13.3) 式的通解由

$$v(t) = (a_1\cos\lambda_1 t + b_1\sin\lambda_1 t)e^1 + \cdots + (a_n\cos\lambda_n t + b_n\sin\lambda_n t)e^n \tag{13.4}$$

给出, 其中常数 a_i 和 b_i 由

$$v(0) = a_1e^1 + \cdots + a_ne^n$$

和

$$\frac{dv}{dt}(0) = \lambda_1b_1e^1 + \cdots + \lambda_nb_ne^n$$

决定.

因而此通解是些独立的振动的叠加. 每个振动的频率由特征值 $\{\lambda_i\}$ 决定. 这就是力学系统 (13.1) 被称作"小振动"的理由.

13.14 小振动 (续)

迄今, 我们一直把线性方程 (13.1) 考虑成为对有限维力学系统的一个逼近. 其基本的理由在于虽然实际的力学系统的解存在, 但难于求出. 我们把 (13.3) 式当作一个好的近似方程.

最终证明, 这个方法具有更为广泛的应用, 甚至对于对那些就连"实际"的力学系统的解的存在性要得以建立都非常困难的无穷维系统的情形, 它也能被应用. 我们用一条两端固定的拉直了的弦线这个力学系统来对此作个解释. 为了解释的简明, 我们假定限制该弦在 xy 平面中运动, 但这对论证绝不是本质性的假定. 我们再假定此弦是均匀的, 两个固定点为 (0,0),(0,1). 那么展示在图 13.10 中的曲线会是我们构形空间中一个可能的构成元素. 在某种意义下, 这个构形是



图 13.10

个"无穷维流形", 并经过适当的构造后, 这个想法可以搞得清楚准确. 然而我们感兴趣的是它在"平衡态"曲线 $C(t) = (t, 0)$ 附近的行为. 对这些曲线我们有 $dx/dt > 0$, 从而用 x 作为独立变量则可描述这些曲线, 也就是给出一个函数 $u(x)$. 换句话说, 我们将大的构形空间换为所有满足 $u(0) = u(1) = 0$ 的一元函数所构成的近似的线性空间 V. 因而 V 被视为我们系统的"切空间", 其意义是说 u 是曲线 $C_s(\cdot)$ 的切向量, 其中的 $C_s(\tau) = (\tau, su(\tau))$.(记住, 我们的构形空间是一个曲线的集合, 故在此构形空间中的曲线是一个曲线的单参数族). 我们现在期望 u 的"动能"为在此弦上所有质点的动能的总和. 在 τ 的质点具有速度 $u(\tau)$ 因而动能为 $\frac{1}{2}mu(\tau)^2$. 如果我们假定质量密度为常量, 则得到动能的表达式

$$K_2(u) = \frac{1}{2}m\int_0^1 u^2 dx.$$

它把 V 做成了在第六章第 6 节所说的准希尔伯特空间.

我们预计势能应依赖于这条弦的伸缩度, 即它是长度的某个函数:

$$\overline{U}(C) = F\left(\int_0^1 \|C'(t)\|dt\right) = F(L),$$

其中 L 为曲线的长, F 为某个光滑函数, 满足 $F(1) = 0$. 对以 x 为参数的曲线, 长度由

$$\int_0^1 \sqrt{1 + \left(\frac{du}{dx}\right)^2} dx.$$

给出. 现有

$$\sqrt{1 + a^2} = 1 + \frac{1}{2}a^2 + \text{高阶项}.$$

因而利用 F 的泰勒展式,

$$F(L) = F'(1)(L-1) + (L-1)\text{的二次项},$$

而

$$L - 1 = \frac{1}{2}\int_0^1 \left(\frac{du}{dx}\right)^2 dx + \int \frac{du}{dx}\text{的高阶项},$$

对于 $\overline{U}$ 的二次项近似的 U_2 为

$$U_2(u) = \frac{C}{2}\int_0^1 \left(\frac{du}{dx}\right)^2 dx.$$

由对 (13.3) 式的类比, 我们预期此弦的"小振动"是方程

$$\frac{d^2 u_t}{dt^2} = -Au_t$$

的解 $u_t(\cdot)$, 其中 A 是个自伴线性算子, 使得

$$(Au, u) = \frac{C}{2}\int_0^1 \left(\frac{du}{dx}\right)^2 dx.$$

现在

$$(Au, u) = \frac{m}{2}\int_0^1 u(Au)dx.$$

由于 $u(0) = u(1) = 0$, 用分部积分我们有

$$\int_0^1 \frac{du}{dx}\cdot\frac{du}{dx} = -\int_0^1 u\left(\frac{d^2u}{dx^2}\right),$$

所以

$$Au = -\frac{C}{m}\frac{d^2u}{dt^2}.$$

因而方程 (13.3) 变为

$$\left(\frac{d}{dt}\right)^2 u_t = \frac{C}{m}\frac{d^2u}{dx^2}, u_t(0) = u_t(1) = 0. \tag{14.1}$$

注意, 我们已经用类比方式的推理导出了 (14.1) 式. 我们既没有详细阐述实际的(非线性) 无穷维力学系统, 也没有任何确保有一个可以解出的这种系统. 另外,

我们实际并不知道函数 F. 我们只需要知道它的形式以及 $F'(1)$ 的值. 尽管如此, 方程 (14.1) 还是给出了对所观察到的物理现象的一个好的解释.

由于具边界条件 $u(0)=u(1)=0$ 的算子 A 是个施图姆 - 刘维尔系统, 故可应用第六章第 6、7 节的结果; 于是解 (14.1) 式正好可以在有限维情形中那样进行. 事实上, 我们可以选取函数

$$u_n(x)=\sin(n\pi x)$$

作为 A 的特征向量的正交基, 其中 u_n 具有特征值 $n^2(c/m)\pi^2$. 因而通解由

$$\begin{aligned}u_t(x)=&(a_1\cos\alpha t+b_1\sin\alpha t)\sin(\pi x)\\&+(a_2\cos 2\alpha t+b_2\sin 2\alpha t)\sin(2\pi x)+\cdots,\end{aligned}$$

其中 $\alpha=(c/m)\pi^2$. 换句话说, 此通解是 "谐波"

$$\sin n\alpha t\sin n\pi x,\quad \cos n\alpha t\sin n\pi x$$

的叠加 (线性组合).

这些也可看作 "驻波", 例如像在图 13.11 中的那样.

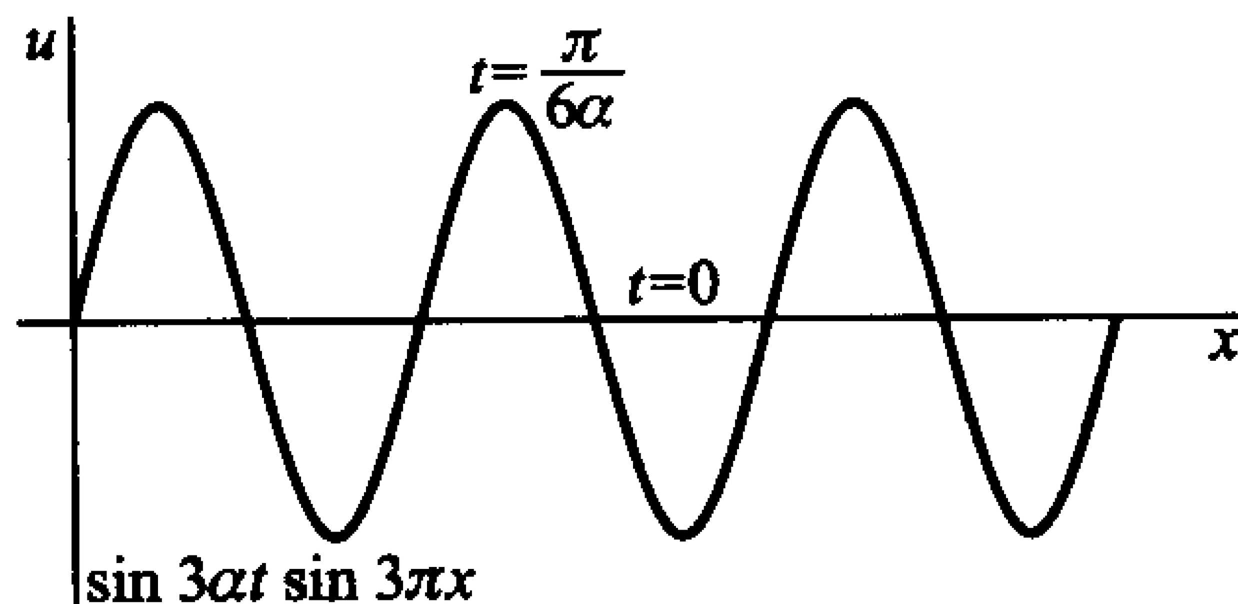


图 13.11

作为这个方法的另一种解释, 我们考虑 n 维的 "振动薄膜". 在这里我们有了 $\mathbb{E}^n$ 中一个具几乎正则边界的区域 D. 我们考虑在 $\mathbb{E}^{n+1}=\mathbb{E}^n\times\mathbb{E}^1$ 中张开的一块膜, 它沿 $\mathbb{E}^n\times\{0\}$ 中的边界 ∂D 被固定. 再一次取 D 上在 ∂D 为 0 的所有函数 u 的空间作为我们对它的构形流形的线性近似 V. 准确地说, 我们让 V 为一个函数空间, 这些函数在 $\overline{D}$ 的某个邻域内是 C^2 类的, 并且在 ∂D 上为零. 因而此薄膜是 $\mathbb{E}^{n+1}$ 中的曲面, 它的点具有形式

$$\prec x^1,\cdots,x^n,u(x^1,\cdots,x^n)\succ,$$

其中 $\prec x^1,\cdots,x^n \succ \in D$. 如同前面那样, 定义动能为

$$K(u)=\frac{1}{2}\int_D u^2,$$

而势能是总体积 (面积) 的某个函数, 它在初始体积 $u(D)$ 为 0. 现在此超曲面的总体积 (面积) 是 (见第十章习题 4.3)

$$\int_D \sqrt{1+\sum\left(\frac{\partial u}{\partial x^i}\right)^2}.$$

因而如前,

$$\overline{U}_2(u)=\frac{C}{2}\int\sum\left(\frac{\partial u}{\partial x^i}\right)^2=\frac{C}{2}D[u,u],$$

其中的 D 是第十二章第 11 节引进的狄利克雷积分. 由格林公式, 我们有 (因为在 ∂D 有 $u=0$)

$$D[u,u]=-\int_D u\Delta u.$$

从而 (13.3) 式的算子 A 由 $-(c/m)\Delta$ 给出, 而方程 (13.3) 成了

$$\frac{d^2}{dt^2}u_i=\frac{c}{m}\Delta u. \tag{14.2}$$

注意, 如果 $n=1$ 那么它就完全是 (14.1) 式. 为了解 (14.2) 式, 我们必须求出 (14.2) 式的特征值的完全组. 如果 $\{u_n\}$ 是这样一组基, 其中 λ_i 为与 u_i 相关联的特征值. 于是 (14.2) 式的通解要由

$$u_t(x)=\sum(a_n\cos\lambda_n t+b_n\sin\lambda_n t)u_n(x).$$

对于证明 $-\Delta$ 具有一个特征向量的完全组, 在 $n>1$ 的情形更困难, 将在习题中予以讨论.

习题

为了研讨特征值问题, 把在 ∂D 为零的 C^1 函数空间换为空间 H_1^D 要方便一些 (我们往回参看第八章第 14 节的习题 4.12 下面部分中关于空间 H_s^D 的定义). 要证明这个置换是合理的, 我们有:

14.1 证明如果 φ 是在 $\overline{D}$ 的一个邻域中为 C^1, 则 $\varphi \in H^D$ 当且仅当 φ 在 ∂D 为零.

14.2 设K 为对光滑函数 f 定义为 $Kf = (1 - \Delta)f$ 的算子, 并如第八章 14 节那样扩张到映射 $H_s \to H_{s-2}$. 设 $g \in H_0$. 由于对任意 $v \in H_1^D$ 有

$$|(g,v)| \leqslant \|g\|_0\|v\|_0 \leqslant \|g\|_0\|v\|_1,$$

试得出存在一个从 H_0^D 到 H_1^D 的有界线性映射, 满足对所有 $v \in H_1^D$ 成立

$$(Lg, v)_1 = (g, v).$$

的结论.

14.3 设f 和 g 在 D 上是个局部可积函数. 这个意思是说对任意测试函数 $\varphi \in C_0^\infty(D), f\varphi$ 和 $g\varphi$ 都是可积的. 我们说微分方程 $Kf = g$ 在 D 上弱成立是指对所有测试函数 $\varphi \in C_0^\infty(D)$ 有 $(f, K_\varphi) = (g, \varphi)$. 证明下面的引理以表明这样的定义推广了解的概念.

引理 14.1　*如果函数 f 和 g 分别在类 $C^2(D)$ 和 $C^0(D)$ 中, 则在弱意义下方程 $Kf = g$ 成立的充要条件是它在经典的意义下成立.*

在证明此引理中你可以假定 h 是 D 上的一个连续函数, 使得 $\int_D h\varphi = 0$ 对所有测试函数 φ 成立, 那么有 $h = 0$.

14.4 证明习题 14.2 中的算子 L 是 K 在弱意义下的一个右逆. 即证明如果 $g \in H_0^D, f = Lg$, 则 $Kf = -g$ 在 D 上弱成立.

14.5 我们现在需要证明如果 g 适当地光滑, 则 f 实际上就在 $C^2(D)$ 中. 粗略地说, 我们需要做的是把 f 和 g 在接近 D 的边界按某种方式“修理圆滑”使得我们可以考虑定义在整个 $\mathbb{R}^n$ 上的调整了的函数, 从而可应用第八章的习题 14.25 和 14.30.

我们的修圆过程只不过是乘以 $C_0^\infty(D)$ 中一个任意却固定的函数 ψ. 作为开始, 证明乘以这样一个 ψ 是一个从 H_s 到自身的有界线性映射, 其中的 s 是任意的.

14.6 我们知道 $D_j = \partial/\partial x_j$ 是个从 H_s 到 H_{s-1} 的有界线性映射, 其中 s 任意. 把这个事实与上面习题的结果结合起来证明, 如果 $Kf = g$ 在 D 上弱成立, 并且如果 ψ 是 C_0^∞ 中任意固定的元素, 则有一个一阶的微分算子 R 定义于整个 $\mathbb{R}^n$, 使得

(1) $K(\psi f) = \psi g + Rg$ 在 D 上弱成立;

(2) $h \mapsto Rh$ 是从 H_s 到 H_{s-1} 的一个有界线性映射, 其中 s 任意.

为了考虑 R 被定义在 $\mathbb{R}^n$ 上, 我们必须以一种显然的方式扩张 ψ 到 $\mathbb{R}^n$. 此证明本质上是个分部积分.

14.7 我们说定义在 D 上的函数 h 是*局部地*在 H_s 中是指对每个 $\varphi \in C_0^\infty(D)$, $\varphi h \in H_s$ (当扩张到 $\mathbb{R}^n$ 时). 用上面的习题证明下面的引理:

引理 14.2 假设 $Kf = g$ 在 D 中弱成立, 而在 D 中局部地 $f \in H_j$, 以及在 D 中局部地 $g \in H_m$, 则在 D 中局部地 $f \in H_{\min\{m+2,j+1\}}$.

[提示: 为了证明这个关键的引理, 首先证明习题 14.6 的弱微分方程对 $\mathbb{R}^n$ 上所有的测试函数 φ 成立. 关键是要存在一个函数 $\chi \in C_0^\infty(D)$ 使得在 ψ 的支集上 $\chi = 1$.(像前面那样) 我们将 χ 扩张到 $\mathbb{R}^n$, 然后对 $\mathbb{R}^n$ 上的每个测试函数 φ, 写为

$$\varphi = \chi\varphi + (1-\chi)\varphi,$$

其中的 $\chi\varphi \in C_0^\infty(D)$. 现在用 $\mathbb{R}^n$ 上测试函数对每个 s 在 H_s 中稠密的性质, 和 $\psi f \in H_j$ 及 $\psi g \in H_m$ 的事实.]

14.8 现在假设 $g \in H_0^D, f = Lg \in H_1^D$(习题 14.2) 及 $kf = g$ 在 D 上弱成立 (习题 14.4). 反复运用上面引理证明如果在 D 中局部地 $g \in H_m$, 则在 D 中局部地 $f \in H_{m+2}$. 由索伯列夫引理得出结论: 如果 $m > n/2 + j$, 且在 D 中局部地 $g \in H_m$, 则 $f = Lg \in C^{j+2}(D)$.

14.9 证明 $\|Lg\|_1 \leqslant \|g\|_0$ 并由第八章的习题 14.31 得出结论说, 如果我们将 L 看作从 H_0^D 到 H_0^D 的算子, 它则是紧的, 并且它的所有特征向量属于 H_1^D. 用习题 14.8 证明每个特征向量属于 $C^\infty(D)$.

13.15 典型变换

在第 1 节到第 5 节中, 我们详细阐述了作为构形空间余切丛上某种类型的流的一个力学系统的概念. 对生成这个流的向量场 X 的定义方程为 $X\rfloor\Omega = -dH$. 因而余切丛的用来挑出这类流的基本性质是 2 - 形式 Ω 的存在性. 最终可以证明, 在研究这些力学方程中有时忘记掉这个流是在余切丛上的而只集中注意形式 Ω 倒是方便的. 例如, 我们可以引进分图表, 虽然它们不是来自构形流形, 但使用它们时向量场 X 具有特别简单的形式. 我们因此要考虑一个流形 N, 它具有一个 2 - 形式 Ω, 并遵从于我们在下面要描述的某些限制. 在这样的流形上, 我们将研究满足 $X\rfloor\Omega = -dH$ 的向量场. 让 H 依赖于时间 t 并同时为 N 上的函数, 故而 X 是个依赖时间的向量场, 这样做是合适的. 这有双重理由. 首先, 它可允许去考虑"非守恒的"力学系统. 其次, 甚至在迄今我们已经介绍过的系统的研究中, 有时做一个与时间有关的坐标变换来简化方程也颇为方便. 它有改变一个与时间无关的向量场为一个与时间相关的向量场的作用. 现在给出定义.

定义 15.1 称流形 N 具有一个哈密顿结构(或者说是一个哈密顿流形) 是说存在一个定义在 N 上的一个外 2 - 形式 Ω, 使得

(i) $d\Omega = 0$, 并且

(ii) Ω 具有极大秩, 这个意思是说 (4.3) 式成立.

注 (a) 如同我们已经知道的, 如果 $N = T^*(M)$, 则 N 是个哈密顿流形, 其中 Ω 由 (4.1) 式给出.

(b) 如果 N 是有限维, 则它必为偶维. 事实上, 条件 (ii) 说限制到每个切空间的 Ω 是个非异的反称双线性形式. 这只能在一个偶维的向量空间上才能发生.

(c) 可以证明, 如果 N 为有限维哈密顿流形, 则总可以找到局部坐标 $q_1, \cdots, q_n, p_1, \cdots, p_n$ 使得

$$\Omega = \sum dp_i \wedge dq_i.$$

[我们知道在 $N = T^*(M)$ 时是这种情形.]这个结果的关键之处 (我们不在这里证明) 是局部地, 所有同样有限维数的哈密顿流形看起来是相像的.

我们现在在 $N \times \mathbb{R}$ 上挑出一类向量场.

定义 15.2 称向量场 $\overline{X}$ 是一个哈密顿向量场是说存在 $N \times \mathbb{R}$ 上一个函数 $H = H_X$, 使得

$$\overline{X}t = \langle \overline{X}, dt \rangle \equiv 1, \tag{15.1}$$

其中 t 是 $\mathbb{R}$ 上的标准坐标, 并看成是 $N \times \mathbb{R}$ 上的一个函数, 并且有

$$\overline{X} \rfloor (\pi^* \Omega - dH \wedge dt) = 0, \tag{15.2}$$

其中 π 为 $N \times \mathbb{R}$ 到 N 上的投影: $\pi(x, t) = x$. 注意, H 被决定到只差 t 的一个函数.

我们令

$$\omega = \pi^* \Omega,$$

故 (15.2) 式可以被写为

$$\overline{X} \rfloor (\omega - dH \wedge dt) = 0. \tag{15.2$'$}$$

如果我们考虑 $N \times \mathbb{R}$ 的切空间的直和分解, 那么条件 (15.1) 式是说我们可以写成

$$\overline{X} = \left(X, \frac{\partial}{\partial t} \right),$$

其中的 X 是 N 上一个与时间有关的向量场；就是说，X 是个规则，它对每个 x 和 t 指定了一个切向量 $X(x,t)\in T_x(N)$. 因为 ω 不含有 dt, 我们可以写为，对任意时间 t 有

$$\overline{X}\rfloor\omega = X(\cdot t)\rfloor\Omega \tag{15.3}$$

[严格地说，此等式应该写成下面的样子：设 $i_t: N\to N\times\mathbb{R}$ 为由 $i_t(x)=(x,t)$ 定义的映射. 于是，

$$i_t^*(X\rfloor\omega) = X(\cdot,t)\rfloor\Omega.] \tag{15.3$'$}$$

另外对任意时间 t 也有

$$\langle\overline{X}, dH\rangle = \langle X(\cdot,t), d\dot{H}(\cdot,t)\rangle + \frac{\partial H}{\partial t}.$$

因而 (15.2) 式可以分成两个方程. 当我们比较不含 dt 的项时，我们得到

$$X(\cdot,t)\rfloor\Omega = -dH(\cdot,t), \qquad \text{对所有固定的}t\text{成立}. \tag{15.2a}$$

于是

$$\overline{X}\rfloor\omega = -dH + \frac{\partial H}{dt}dt. \tag{15.2b}$$

注意 (15.2a) 式，它恰好是在第 5 节所陈述条件的创新，即 H(因而 X) 现在可以依赖于时间.

定义 15.3 称一个微分同胚 $\bar{\varphi}: N\times\mathbb{R}\to N\times\mathbb{R}$ 为一个典型变换是说

(i) $\bar{\varphi}^*(\omega) = \omega - dW\wedge dt$, 其中 $W = W_{\bar{\varphi}}$ 是某个依赖于 $\bar{\varphi}$ 的函数，并且

(ii) $\bar{\varphi}$ 是保持时间不变的，即 $\bar{\varphi}$ 具有形式 $\bar{\varphi}(x,t) = (\varphi(x,t),t)$, 其中 $\varphi(\cdot,t)$ 对每个 t 为 N 的微分同胚.

注意，如果 $\bar{\varphi}$ 是个典型变换，则 $\bar{\varphi}^{-1}$ 也是，并且

$$W_{\bar{\varphi}^{-1}} = -(\bar{\varphi}^{-1})^* W_{\bar{\varphi}}. \tag{15.4}$$

我们也看出，如果 $\bar{\varphi}$ 和 $\bar{\psi}$ 为典型变换，那么 $\bar{\psi}\circ\bar{\varphi}$ 也是，其中

$$W_{\bar{\psi}\circ\bar{\varphi}} = \bar{\psi}^* W_{\bar{\varphi}} + W_{\bar{\psi}}. \tag{15.5}$$

这些事实都直接从定义得到，当作习题留给读者.

我们下一步所注意到的是, 如果 $\bar{\varphi}$ 是个典型变换而 $\overline{X}$ 是个哈密顿向量场, 则 $\overline{\varphi}^*(\overline{X})$ 也是个哈密顿向量场. 事实上,

$$\begin{aligned}&\bar{\varphi}^*\overline{X}\rfloor[\omega - d(W_{\bar{\varphi}} + \bar{\varphi}^* H)\wedge dt]\\&=\bar{\varphi}^*\overline{X}\rfloor[\bar{\varphi}^*\omega - \bar{\varphi}^*(dH\wedge dt)]\\&=\bar{\varphi}^*[X\rfloor(\omega - dH\wedge dt)] = 0.\end{aligned}$$

因而我们可取 $H_{\varphi^* X}$ 为

$$H_{\bar{\varphi}^* X} = W_{\bar{\varphi}} + \bar{\varphi}^* H \tag{15.6}$$

设 $\overline{X}$ 为一个哈密顿向量场, $\overline{\varphi}$ 为 $N\times\mathbb{R}\to N\times\mathbb{R}$ 的映射, 它是由按 $\overline{X}$ 所生成的流让此系统从 $t=0$ 演变得到的. 就是说, 设 $\varphi(\cdot,t)$ 定义为使曲线 $t\mapsto\varphi(x,t)$ 是 (与时间相关的) 向量场 X 的解曲线, 它在 $t=0$ 时通过 x. 换一种方式说, 即曲线 $t\mapsto(\varphi(x,t),t)$ 是向量场 $\overline{X}$ 的解曲线, 它在时间 0 通过 $(x,0)$(见图 13.12 和图 13.13).

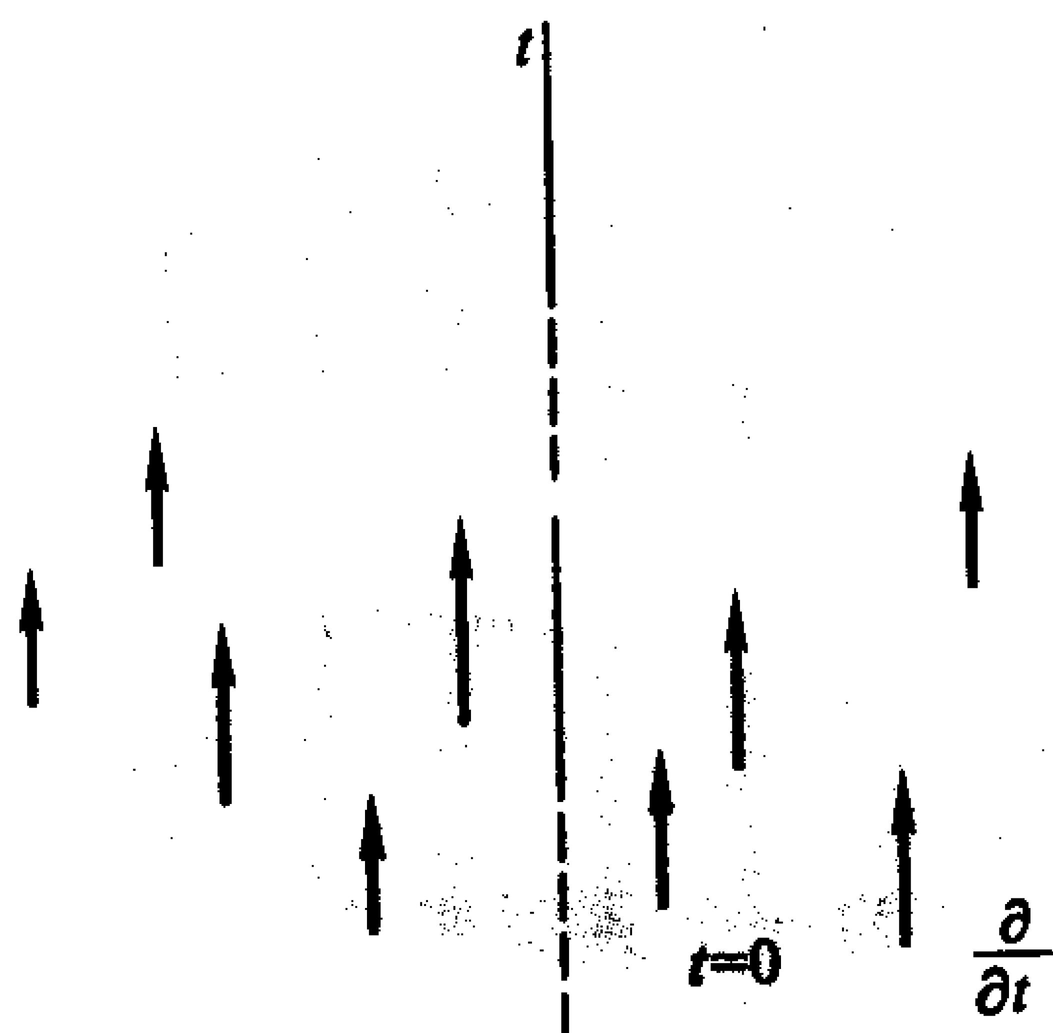


图 13.12

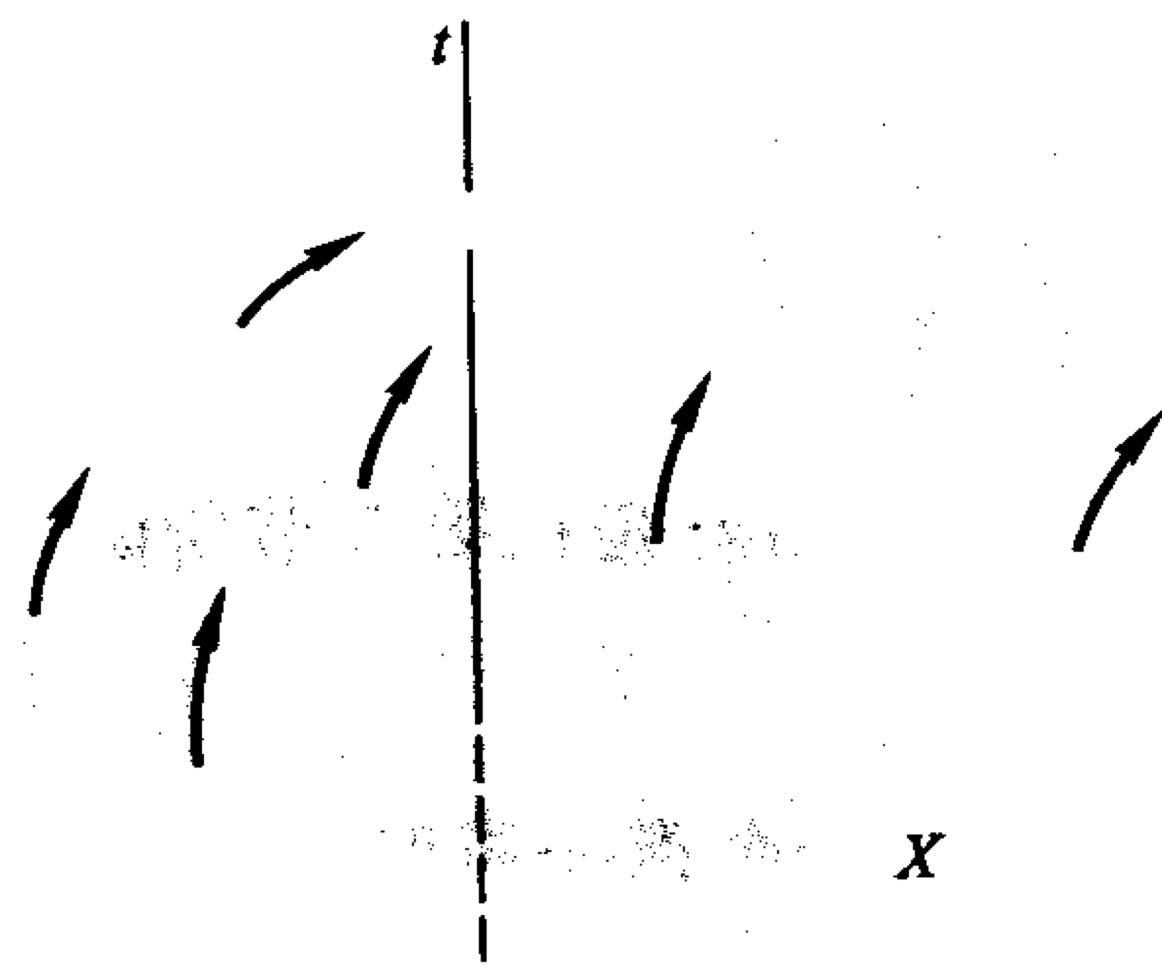


图 13.13

注意, 由 $\bar{\varphi}$ 的定义得到

$$\bar{\varphi}_*\left(\frac{\partial}{\partial t}\right) = X_{\varphi(x,t)}.$$

我们断定 $\bar{\varphi}$ 是个典型变换. 事实上, 因为 $\varphi\circ i_t(x,t) = (\varphi(x,t),t)$, 故

$$i_t^*(\bar{\varphi}^*\omega) = (\bar{\varphi}\circ i_t)^*\pi^*\Omega = \varphi(\cdot,t)^*\Omega.$$

但是, 由 (15.2a) 式,

$$\frac{d}{ds}(\varphi(\cdot,s)^*\Omega) = \varphi(\cdot,s)^* D_{X(\cdot,t)}\Omega = 0.$$

因此

$$i_t^*\bar{\varphi}^*\omega = \Omega,$$

或者, 换句话说, $\varphi^*\omega$ 具有形式 $\omega + \theta \wedge dt$. 只要取与 $\partial/\partial t$ 的内部乘积就可以决定 θ, 这是因为 ω 不含 dt 的缘故. 但是

$$\left(\frac{\partial}{\partial t}\rfloor\bar{\varphi}^*\pi^*\Omega\right) = \bar{\varphi}^*\left[\bar{\varphi}^*\left(\frac{\partial}{\partial t}\right)\rfloor\pi^*\Omega\right] = \bar{\varphi}^*(X\rfloor\pi^*\Omega)$$
$$= \bar{\varphi}^*\left(-dH + \frac{\partial H}{\partial t}dt\right),$$

故 $\theta = \bar{\varphi}^*dH$. 因而 $\bar{\varphi}$ 是个典型变换, 并且

$$W_{\overline{\varphi}} = -\bar{\varphi}^*H_X. \tag{15.7}$$

注意, (15.7) 式正是我们想从 (15.6) 式得到的东西. 事实上, $\varphi^*X = \partial/\partial t$, 从而我们可取 $H_{\partial/\partial t} = 0$.

等式 (15.6) 和 (15.7) 可以以下面的方式结合起来使用: 假设 $H = H_0 + H_1$, 而我们已经知道了该如何解对应于 H_0 的微分方程. 换句话说, 我们能够求出对应于向量场 X_0 的映射 $\bar{\varphi}$, 其中 $X_0\rfloor(\omega - dH_0 \wedge dt) = 0$. 如果

$$X\rfloor(\omega - dH \wedge dt) = 0,$$

则由 (15.6) 式 φ^*X 是一个其对应哈密顿函数为 $\bar{\varphi}^*H_1$ 的向量场.

这个方法首先由拉格朗日在研究 n 体问题时引进的. 我们可以设 H_0 为忽略掉势能中来自行星作用的项而得到哈密顿函数, 而让 H_1 为哈密顿 H 的剩余部分. 对 H_0 的解于是按开普勒定律由行星绕太阳运动给出. 为了使讨论简明, 我们只关注运动是椭圆的那种相空间. 于是行星的运动由给出各个椭圆的种种参数而被完全确定 (诸如椭圆所在平面, 其长轴, 离心率, 等等) 并且告诉我们在时刻 $t = 0$ 时行星在它的椭圆轨道上的位置. 这对应于映射 φ 的运用. 于是把整个系统的运动方程看为对每个椭圆的种种参数的微分方程. 这对应于研究向量场 φ^*X. 引进椭圆的参数为 "广义的" 坐标的思想是引向流形上不变量计算概念的关键步骤之一.

我们已经知道, 解对应于一个哈密顿 H 的微分方程和寻找一个满足 (15.7) 式的映射 $\bar{\varphi}$ 是同一回事. 在某种条件下, 它可以化为求某个偏微分方程的解. 假设我们有局部坐标 $q_1, \cdots, q_n, p_1, \cdots, p_n$, 使得 $\Omega = \sum dp_i \wedge dq_i$. 设 $V =$

$V(q_1,\cdots,q_n,p_1,\cdots,p_n,t)$ 为一个使得映射 φ_1,φ_2 为微分同胚的函数, 其中

$$\varphi_1(q_1,\cdots,q_n,p_1,\cdots,p_n,t)=\prec q_1,\cdots,q_n,\frac{\partial V}{\partial q_1},\cdots,\frac{\partial V}{\partial q_n},t\succ,$$
$$\varphi_2(q_1,\cdots,q_n,p_1,\cdots,p_n,t)=\prec \frac{\partial V}{\partial p_1},\cdots,\frac{\partial V}{\partial p_n},p_1,\cdots,p_n,t\succ.$$

我们断定 $\bar{\varphi}=\varphi_1\circ\varphi_2^{-1}$ 是个典型变换, 并且

$$W_{\bar{\varphi}}=\varphi_2^{-1*}\frac{\partial V}{\partial t}. \tag{15.8}$$

注意有 $\omega=d(\sum p_i dq_i)=-d(\sum q_i dp_i)$. 因而

$$\begin{aligned}
\varphi^*\omega-\omega&=d(\bar{\varphi}^*\sum p_i dq_i+\sum q_i dp_i)\\
&=d\varphi_2^{-1*}(\varphi_1^*\sum p_i dq_i+\varphi_2^*\sum q_i dp_i)\\
&=d\varphi_2^{-1*}\left(dV-\frac{\partial V}{\partial t}dt\right)\\
&=\varphi_2^{-1*}ddV-d\varphi_2^{-1*}\frac{\partial V}{\partial t}\wedge dt\\
&=-d\left(\varphi_2^{-1*}\frac{\partial V}{\partial t}\right)\wedge dt.
\end{aligned}$$

如果将它代入 (15.7) 式, 我们则看出 φ 为我们的微分方程的解当且仅当

$$\varphi_2^{-1*}\frac{\partial V}{\partial t}+\bar{\varphi}^*H=0.$$

但是 $\bar{\varphi}^*H=\varphi_2^{-1*}\varphi_1^*H$, 所以我们可以把这个方程写为

$$\frac{\partial V}{\partial t}+\varphi_1^*H=0 \tag{15.9}$$

或者

$$\frac{\partial V}{\partial t}+H\left(q_1,\cdots,q_n,\frac{\partial V}{\partial q_1},\cdots,\frac{\partial V}{\partial q_n},t\right)=0. \tag{15.9$'$}$$

方程 (15.9) 以哈密顿 - 雅可比方程而知名.

因此我们有了一个 (局部地) 解运动方程的方案: 求 (15.9) 式的一个解 V, 而它具有使 φ_1,φ_2 为微分同胚的性质. 在某些条件下, 适当的坐标选取可使我们用"分离变量"法解出 (15.9) 式. 在下列例题中我们要解释此方法.

例题 15.1 再谈中心力运动问题. 这里当使用平面极坐标时有

$$H=\frac{1}{2m}\left(p_r^2+\frac{p_\theta^2}{r^2}\right)+U(r).$$

方程 (15.9) 成为

$$\frac{\partial V}{\partial t}+\frac{1}{2m}\left[\left(\frac{\partial V}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial V}{\partial\theta}\right)^2\right]+U(r)=0.$$

由于变量 t 和 θ 不是显式地出现在该方程中, 我们要求一个形如

$$V=V_1(t)+V_2(\theta)+V_3(r)$$

的解, 并得出结论说 $V_1'(t)$ 和 $V_2'(\theta)$ 只依赖于 r, 从而必为常数. 我们可以由此写出

$$V_1'(t)=-E,\quad V_2'(\theta)=A_\theta,$$

其中 E 和 A_θ 为常数. (这反映了能量守恒和角动量守恒). 我们于是得到方程

$$\frac{\partial V}{\partial r}=V_3'(r)=\sqrt{2m(E-U(r))-\frac{A_\theta^2}{r^2}}.$$

因而

$$V=A_\theta\theta+\int_0^r\left[2m(E-U(s))-\frac{A_\theta^2}{s^2}\right]^{1/2}ds-Et$$

是 (15.9) 式的一个解. 在这里我们把 V 想成是变量 r,θ,E,A_θ 的函数. 那么, 映射 φ_2 由

$$\varphi_2(r,\theta;E,A_\theta)=\prec -t+\int_0^r\frac{mds}{\left[2m(E-U(s))-\frac{A_\theta^2}{s^2}\right]^{1/2}},$$

$$\theta-\int_0^r\frac{A_\theta ds}{s^2\left[2m(E-U(s))-\frac{A_\theta^2}{s^2}\right]^{1/2}},E,A_\theta\succ.$$

现在映射 $\varphi_1\circ\varphi_2^{-1}$ 把这个流变为一个常值流, 故我们必有

$$t-t_0=\int_0^r\frac{mds}{\left[2m(E-U(s))-\frac{A_\theta^2}{s^2}\right]^{1/2}}$$

和

$$\theta-\theta_0=\int_0^r \frac{A_\theta ds}{s^2\left[2m(E-U(s))-\frac{A_\theta^2}{s^2}\right]^{1/2}},$$

其中 t_0 和 θ_0 为常数. 注意这些等式的第二个, 它给出了显式表达的轨道, 我们于是可以解出它从而给出 t 的函数 r.

例题 15.2 简谐振子. 在这里

$$H=\frac{p^2}{2m}+\frac{kq^2}{2},$$

故方程 (15.9) 成为

$$\frac{1}{2m}\left(\frac{\partial V}{\partial q}\right)^2+\frac{kq^2}{2}+\frac{\partial V}{\partial t}=0.$$

再次因为时间并没有显式出现, 故可写出

$$V=-Et+W,$$

其中 W 仅是 q 的函数, 它必定满足

$$\frac{1}{2m}(W')^2+\frac{kq^2}{2}=E,$$

或者

$$W=\sqrt{mk}\int_0^q\left(\frac{2E}{k}-s^2\right)^{1/2}ds.$$

于是

$$\frac{\partial V}{\partial E}=\left(\frac{m}{k}\right)^{1/2}\int_0^q\left(\frac{2E}{k}-s^2\right)^{-1/2}ds-t.$$

从而有

$$\begin{aligned}t-t_0&=-\left(\frac{m}{k}\right)^{1/2}\int_0^q\frac{ds}{\left(\frac{2E}{k}-s^2\right)^{1/2}}\\&=-\left(\frac{m}{k}\right)^{1/2}\arccos\left(\frac{k}{2E}\right)^{1/2}q.\end{aligned}$$

解出用 t 表示的 q, 给出

$$q=\sqrt{\frac{2E}{k}}\cos\left(\frac{k}{m}\right)^{1/2}(t-t_0).$$

例题 15.3 由两个固定点的质量所吸引的质点运动. 这里的

$$H = \frac{1}{2m}(P_x^2 + P_y^2 + P_z^2) + \frac{A}{r_1} + \frac{B}{r_2},$$

其中 r_1 和 r_2 分别为到这两个点的距离, 而 A 和 B 为常数 (由这两个点的质量所决定).

为了解决这个问题, 引进所谓的椭圆坐标较为合适. 我们假定这两个固定点位于 x 轴上, 每个与原点的距离都为 c. 我们可取 $c = 1$ 以使计算简明 (见图 13.14). 在 xy 平面中定义局部坐标 ξ 和 η 为

$$\xi = \frac{1}{2}(r_1 + r_2), \quad \eta = \frac{1}{2}(r_1 - r_2).$$

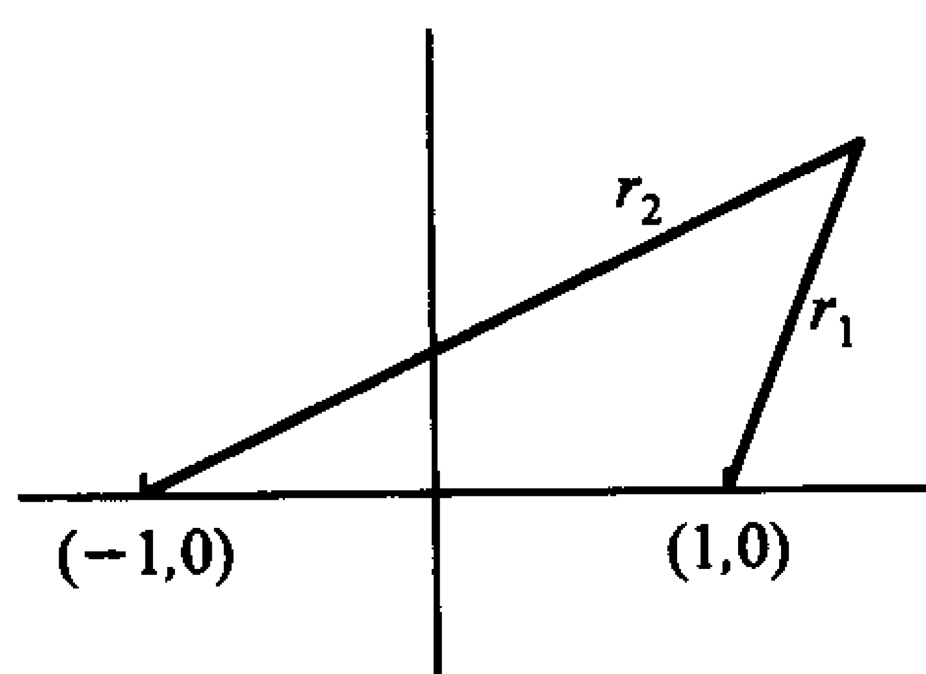


图 13.14

因此曲线 $\xi=$ 常数代表了椭圆, 其半长轴为 ξ, 焦点在两个固定点, 而 $\eta=$ 常数则为双曲线, 半长轴为 η, 并有相同的焦点. 注意, 这时有 $0 < |\eta| \leqslant 1 \leqslant |\xi| < \infty$. 这两条曲线的方程分别为

$$\frac{x^2}{\xi^2} + \frac{y^2}{\xi^2 - 1} = 1 \quad 和 \quad \frac{x^2}{\eta^2} - \frac{y^2}{1 - \eta^2} = 1,$$

所以

$$x^2 = \xi^2\eta^2 \quad 和 \quad y^2 = (\xi^2 - 1)(1 - \eta^2).$$

于是

$$\frac{dx}{x} = \frac{d\xi}{\xi} + \frac{d\eta}{\eta}, \quad \frac{dy}{y} = \frac{\xi d\xi}{\xi^2 - 1} - \frac{\eta d\eta}{1 - \eta^2},$$

因此

$$dx^2 + dy^2 = (\xi^2 - \eta^2)\left(\frac{d\xi^2}{\xi^2 - 1} + \frac{d\eta^2}{1 - \eta^2}\right).$$

如果我们现在绕 x 轴旋转得到空间中可类比的柱坐标, 我们便有空间坐标

$$\prec x, \rho, \theta \succ \quad \text{和} \quad \prec \xi, \eta, \theta \succ,$$

而欧几里得度量由

$$\begin{aligned} dx^2 + dy^2 + dz^2 &= dx^2 + d\rho^2 + \rho^2 d\theta^2 \\ &= (\xi^2 - \eta^2)\left(\frac{d\xi^2}{\xi^2 - 1} + \frac{d\eta^2}{1 - \eta^2}\right) + (\xi^2 - 1)(1 - \eta^2) d\theta^2 \end{aligned}$$

给出, 同时也有

$$U = \frac{A}{r_1} + \frac{B}{r_2} = \frac{Ar_2 + Br_1}{r_1 r_2} = \frac{1}{\xi^2 - \eta^2}(\alpha\xi + \beta\eta),$$

其中 $\alpha = A + B, \beta = A - B$. 于是 H 有形式

$$\begin{aligned} H(\xi, \eta, \theta, P_\xi, P_\eta, P_\theta, t) = \frac{1}{2m(\xi^2 - \eta^2)} \times \bigg[&(\xi^2 - 1) P_\xi^2 + (\eta^2 - 1) P_\eta^2 \\ &+ \left(\frac{1}{\xi^2 - 1} + \frac{1}{1 - \eta^2}\right) P_\theta^2 + \alpha\xi + \beta\eta \bigg]. \end{aligned}$$

由于 t, θ 在 (15.9) 式中没有显式出现, 我们可以写出

$$V = -Et + A_\theta \theta + W,$$

其中 W 必定满足

$$\begin{aligned} (\xi^2 - 1)\left(\frac{\partial W}{\partial \xi}\right)^2 + (\eta^2 - 1)\left(\frac{\partial W}{\partial \eta}\right)^2 + \left(\frac{1}{\xi^2 - 1} + \frac{1}{1 - \eta^2}\right) A_\theta^2 + \alpha\xi + \beta\eta \\ = 2m(\xi^2 - \eta^2) E. \end{aligned}$$

注意, 如果令 $W = W_1(\xi) + W_2(\eta)$, 这个方程则被分成两个, 我们则可用求积分的办法显式解出其中每一个. 这给出了原来的运动方程的解. 详细做法留给读者去完成.

参考文献

第一、二、七章

在任意系数域上的更广泛的线性代数方面的工作可参考:

BIRKHOFF, G., and S. MACLANE, *A Survey of Modern Algebra*, 3rd ed., Macmillan, New York, 1965.

HOFFMAN, K., and R. KUNZE, *Linear Algebra*, Prentice-Hall, Englewood Cliffs, N.J., 1961.

对交换环上 (而不是域) 的线性代数可见:

LANG, S., *Algebra*, Addison-Wesley, Reading, Mass., 1965.

ZARISKI, O., and P. SAMUEL, *Commutative Algebra*, 2 vols., Van Nostrand, Princeton, 1959, 1960.

第三章

困难但值得一读的是:

DIEUDONNÉ, J., *Foundations of Modern Analysis*, Academic Press, New York, 1960.

此书中的微分记号是 $Df(x)$ (而非 dF_α) 和 $Df\alpha)\cdot\xi$ (而非 $dF_\alpha(\xi)$).

第四、五章

一般 (点集) 拓扑的好书是:

KELLEY, J., *General Topology*, Van Nostrand, Princeton, 1961.

SIMMONS, G., *Topology and Modern Analysis*, McGraw-Hill, New York, 1963.

许多没有讲的巴拿赫空间和希尔伯特空间的标准例子需要勒贝格积分的知识, 因此超出了我们的范围. 但是, 感兴趣的读者可以在 Simmons 和诸如下面的书中继续深入到巴拿赫和希尔伯特空间的抽象理论中.

HILLE, E., and R. S. PHILLIPS, *Functional Analysis and Semigroups*, American Mathematical Society, Providence, R. I., 1957.

MURRAY, F. J., *An Introduction to Linear Transformations in Hilbert Space*, Princeton University Press, Princeton, 1941.

RIESZ, F., and B. Sz-NAGY, *Functional Analysis*, Ungar, New York, 1955.

YOSIDA, K., *Functional Analysis*, Springer, Berlin, 1965.

Murray 和 Yosida 的书更深更难一些.

第六章

关于常微分方程的标准书是:

BIRKHOFF, G., and G. C. ROTA, *Ordinary Differential Equations*, Ginn, Boston, 1962.

HUREWICZ, W., *Lectures on Ordinary Differential Equations*, M. I. T. Press, Cambridge, Mass., 1958.

高等的专著是:

CODDINGTON, E., and N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

HARTMAN, P., *Ordinary Differential Equations*, Wiley, New York, 1964.

第八章

这一章致力于容度理论. 对现代数学而言需要更强的勒贝格测度和积分的理论. 对于一维方面的理论读者可参看:

RUDIN, W., *Principles of Mathematical Analysis*, 2nd ed., McGraw-Hill, New York, 1964.

对一般理论读者可参看:

HALMOS, P., *Measure Theory*, Van Nostrand, Princeton, 1961.

在第十二章末尾的习题汇编中讨论了包含"留数计算"在内的计算某些定积分的其他方法,读者可在第八、十一章之后独立地阅读它们.

对于积分与"广义函数"的关系可见:

GEL'FAND, I. M., and G. E. SHILOV, *Generalized Functions*, vol. 1, Academic Press, New York, 1964.

读此书需要一点复变函数论的知识.

第九、十、十一章

更抽象的处理可见:

LANG, S., *Introduction to Differentiable Manifolds*, Interscience, New York, 1962.

不太抽象的处理可见:

FLANDERS, H., *Differential Forms with Applications to Physical Sciences*, Academic Press, New York, 1963.

FLEMING, W., *Functions of Several Variables*, Addison-Wesley, Reading, Mass., 1965.

SPIVAK, M., *Calculus on Manifolds*, Benjamin, New York, 1965.

对更广泛的微分几何的讲述可见:

WILLMORE, T., *An Introduction to Differential Geometry*, Oxford University Press, London, 1959.

第十二章

位势理论的经典著作是:

KELLOGG, O., *Foundations of Potential Theory*, Springer, Berlin, 1929.

关于平面上调和函数与解析函数之间的关系可在有关后面这个学科的标准教科书中找到,诸如:

AHLFORS, L., *Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1966.

HILLE, E., *Analytic Function Theory*, Ginn, Boston, 1959.

第十三章

力学方面的一本标准的现代著作是:

GOLDSTEIN, H., *Classical Mechanics*, Addison-Wesley, Reading, Mass., 1952.

对经典的面向天文学方面的力学可见:

POINCARÉ, H., *Leçons de mécanique céleste*, Gauthier-Villars, Paris, 1905~1910.

WHITTAKER, E. T., *Analytical Dynamics*, 4th ed., Cambridge University Press, Cambridge, England, 1937.

对经典力学的一种休闲式的几何研习可见:

LANCZOS, C., *The Variational Principles of Mechanics*, University of Toronto Press, Toronto, 1960.

以本章的精神来处理经典力学的一本书, 它同时研究了统计力学和量子力学, 见:

MACKEY, G., *Mathematical Foundations of Quantum Mechanics*, Benjamin, New York, 1965.

对变分法的几何应用方面一种漂亮的处理方法可见:

MILNOR, J., *Morse Theory*, Princeton University Press, Princeton, 1965.

记　号

一般约定

□　证明结束

$\mathbb{R}$　实数系

$\mathbb{Z}$　整数

$\mathbb{C}$　复数系

$\mathbb{E}^n$　n 维欧几里得空间

对 n 联 (元) 组的黑体字母：$\boldsymbol{\alpha} :\prec \alpha_1, \cdots, \alpha_n \succ$

特殊符号

符号	页数	符号	页数
$(\forall x)$	2	$\Leftrightarrow$	6
$(\exists x)$	2	$\in$	7
&	4	$\subset$	7
$\sim$	4	$\varnothing$	8
$\Rightarrow$	4	$\{\ \}$	8

索 引

D

J

K

R

S

T

W